Designing for Socially Acceptable Security Technologies

PhD Thesis

University College London

2014

Timothy Greig Nissen

Declaration:

I, Timothy Greig Nissen confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Acknowledgements

Despite the solitary nature of undertaking a PhD it is not the product of a vacuum, rather it is a social experience. As such there are people to whom I am grateful; to whom I owe debts of thanks. These are the people I acknowledge here.

First and foremost to my supervisors, Brian Balmer and Alex Braithwaite. With hand on heart I thank you both for your unwavering assistance, guidance, encouragement, and support. If either of you were ever in doubt as to the potential of this dissertation actually being brought to fruition, you were both masters of the poker-face.

To my fellow cohort and the staff on the SECReT programme in the Department for Security and Crime Science, thank you for an amazing four years. Special thanks to my MRes supervisors, Kevin Chetty and Kate Bowers. Also to Hérve Borrion and Noemie Bouhana for their friendship, good travel company, and the sleepless nights resulting from a beyond-ambitious, never-ending stream of research funding bids.

To all those within the Department for Science and Technology Studies, and especially Inga Kroener, thank you for taking me in and providing me with a second home here at UCL. Also for encouraging me to participate fully within the STS world despite the sometimes alien nature of my research project.

To those individuals and companies within the field of security technology research and design who agreed to be interviewed and who gave me access to their work spaces on the proviso I respect and maintain their anonymity, know that I have maintained my part of the bargain. Without your participation and assistance this research project would not have been possible, so for that and for affording me your trust, I owe you all an immense debt of gratitude.

To my family, thank you for producing a son and brother with the capacity and courage to question everything I was ever taught to be true. To my friends, thank you for your understanding, toleration, and inebriation when I ranted about security, ethics, function-creep, state overreach, etc., etc., like a foil-hat-wearing nutter.

And finally to my wife Kerrie, there simply aren't the words to express everything that I need to tell you, and so 'thank you and I love you' will have to suffice.

Abstract

Security technologies (STs) are increasingly being positioned, developed, and implemented as technological-fixes for addressing crime; never more so than in the wake of the numerous terrorist attacks beginning with September 11th 2001. However, despite the purported security benefits afforded citizens by these technologies, their smooth assimilation into society is never assured. STs which evoke social controversy and resistance fail to survive unscathed over the mid- to long-term; subjected instead to enforced modification, restrictions on acquisition, restrictions on use, or in the worst case scenario - outright banning. Such controversies can negatively affect the companies designing these STs, end-users who employ them, governments who authorise them, and citizens whose security may genuinely remain compromised.

The aim of this thesis is to assist the developers and designers of STs in anticipating and mitigating negative societal responses to their technologies upstream in the design process. The logic being that; by targeting STs before they are completed those elements of design most likely to evoke controversy can be modified, which in turn will produce STs the public are more likely to afford legitimacy through acceptance.

To achieve this aim, three objectives were set. The first was to identify the causes of social controversies arising from the design and operation of STs. Through repeated focussed case-studies of previous controversial STs a taxonomy of forty-three commonalities of controversy was produced. The second goal was to generate guidelines for the development of future methodological design-tools that could be produced to assist those developing STs in identifying these controversies. This was achieved by conducting interviews with scientists and engineers actively involved in the design and production of STs. Finally, this taxonomy and guidelines were applied to produce two prototypes of potential design tools; with one subsequently applied to an ongoing ST design project.

Table of Contents

List of Abbreviations

9/11	September 11th 2011
ACPO	Association of Chief Police Officers
ANPR	Automated Number-Plate Recognition
ASBO	Anti-Social Behaviour Order
ATD	Automated Threat Detection
ATE	Anticipatory Technology Ethics
BXS	Backscatter X-ray Scanner
CCTV	Closed Circuit Television
CoE	Council of Europe
DeSAST	Designing for Socially Acceptable Security Technologies
DfT	UK Department for Transport
DHS	US Department of Homeland Security
DNA	Deoxyribonucleic Acid
DPA	UK Data Protection Act 1998
DPTAC	Disabled Persons Transport Advisory Committee
ECHR	European Convention on Human Rights
ECtHR	European Court of Human Rights
EU	European Union
FCC	Framework for Common Controversies within security technologies
HRA	UK Human Rights Act 1998
IA	Impact Assessment
ICA	UK Identity Card Act 2006

ICRP	International Commission on Radiological Protection
ICT	Information and Communications Technology
ID	Identity
LLW	Less Lethal Weapon
LSE	London School of Economics and Political Science
MCDM	Multi-Criteria Decision Making
MWS	Millimetre-Wave Scanner
NATO	North Atlantic Treaty Organisation
NCRP	National Council on Radiation Protection and Measurements
NDNAD	UK National DNA Database
NGO	Non-Governmental Organisation
NIR	National Identity Register
NIRNo	National Identity Registration Number
NIS	National Identity Scheme
PCA	UK Protection of Children Act 1978
PET	Privacy Enhancing Technology
RFID	Radio Frequency Identification
RIPA	UK Regulation of Investigatory Powers Act 2000
ST	Security Technology
STEM	Science, Technology, Engineering, and Mathematics
TA	Technology Assessment
TSA	US Transportation Security Administration
UK	United Kingdom
UNCRC	United Nations Convention on the Rights of the Child
US	United States of America

VHF	Very High Frequency
WBS	Whole Body Scanners

List of Figures

List of Tables

Preface

My motivations for undertaking this dissertation were a combination of exasperation and dismay over the development and deployment of numerous security-centric technologies supposedly justified by the threat of terrorism in our post 9/11 World. In particular I was struck by just how poorly conceived, poorly implemented, and disproportionate a number of these interventions ultimately turned out to be. I was also struck by what seemed an inability by those commissioning, designing, developing, and deploying whatever the latest *social unacceptable security technology* was to the learning of lessons from their previous failure(s). It is worth stating here that these assertions should in no way be read as a personal rejection of a government's need to employ technologies in their efforts to reduce crime and afford their citizens a measure of security. Despite my unashamedly libertarian inclinations I fully acknowledge the need for, and value of, proportionate and justifiable technologies designed to achieve these goals. I simply assert that socially unacceptable security technologies are not the way forward, and ultimately bring harm to governments, end-users, designers, and society.

And so my dissertation, *Designing for Socially Acceptable Security Technologies*, focusses on the upstream development of security technologies in an effort to assist designers identify and mitigate potential sources of social controversy arising from their design choices *before* their technologies are deployed. It is about enabling hard-scientists, engineers, and mathematicians to recognise and incorporate ethical, social, and rights-based concerns *in their technical endeavours*. However, while my dissertation seeks to address the problem of socially unacceptable security technologies, the inherent complexity of producing measures for addressing crime and national security means there are no silver bullets in this area. My efforts focus on only one element of a much wider process.

What will strike the reader of my dissertation is its unconventional approach and format, and yet once one takes into account the following three factors which underlie this research project, this unconventionality is explainable. Firstly I am a member of the initial cohort of PhD candidates within UCL's Security Science Doctoral Training

Centre (SECReT), situated within the Department of Security and Crime Science. SECReT pursues research into the crime and security domains by adopting the widest multidisciplinary approach possible; bringing together all natural and social sciences, engineering, mathematics, as well as law and ethics. This multidisciplinarity is not used as a catchphrase; research projects are expected to cross domains and candidates are required to find their two supervisors from different departments. As my dissertation is one of the first from within this newly-constructed multidisciplinary Centre, the 'typical' SECReT PhD format has not yet been created.

Secondly there is the expectation on producing tangible scientific and technical output. This is perhaps unsurprising given; (a) we are situated within the Engineering Faculty, (b) the multidisciplinary focus forces the interaction of social sciences and STEM fields, (c) the crime/security problems we are tackling are not abstract notions but ongoing concrete problems, and (d) our PhDs are funded by the Engineering and Physical Sciences Research Council (the UK's main agency for funding research in engineering and the physical sciences). For those with a technical background this is less of a challenge; however my primary academic background is in law, human rights and ethics. This legal knowledge has also been supplemented with a developing level of expertise in the discipline of Science and Technology Studies (STS) (specifically within the sociology of science and technology) as a result of my teaching within this field and the supervision of Professor Brian Balmer; one of my PhD supervisors. While law and STS have value to add to SECReT, on a personal level this skill-set does not afford me the required knowledge to undertake cutting-edge natural science or engineering research. However, what it does provide me with is the knowledge-base, understanding, and outsider's standpoint from which to assess the same problems being tackled by those scientists and engineers, but through a different lens. I have therefore sought to incorporate elements of law, STS, and security and crime science in an effort to facilitate the production of something practical and concrete – in this case methodological tools to assist the developers of security technologies to identify and incorporate social, ethical, and legal concerns into their designs *before* their products are released.

Thirdly, my dissertation does not follow a standard 'aim-hypothesis-method-results' format; a format which works well when the researcher is seeking to address a small gap in knowledge of an area which has been extensively researched. This is because my chosen area of research (the holistic examination of the design of security technologies and their associated social, legal, and ethical problems) is not developed enough to have created small identifiable gaps in the remainder. While there exists incredibly detailed, extensive research into both individual security technologies/problems (i.e. CCTV, ID cards, privacy, etc.) and categories of security technologies/problems (i.e. less-lethal weapons, digital technologies, surveillance technologies, data protection, etc.), relatively less research has been conducted where the focus is universal (looking across all of these ostensibly disparate technologies and associated problems to identify the possible presence of commonalities), let alone attempting to apply this knowledge to produce methodological design tools. As a direct result of this broad focus and the lack of an existing corpus of literature, when expanding upon the *aim* of my dissertation in Chapter 1 I have chosen to frame the constituent components as three *research objectives* rather than the more conventional *research questions*. I feel this 'objectives' depiction is a less artificial in the overall context of my research project. The three research objectives (1. identifying commonalities of controversy which exist across different security technologies; 2. generating guidelines for the development of future design tools to assist those developing STs in addressing these social controversies; and 3. applying these guidelines to begin the process of creating prototype methodological design tools) are all diverse in nature. This has the effect of dividing my research project into distinct stages; which further precludes the aim-method-results format.

Throughout this entire undertaking I have apprehensively struggled with the unorthodox nature of my research project. However, upon completion I am proud of this lack of orthodoxy. For I contend that while the physical depiction of knowledge-creation (i.e., a written dissertation or research paper) may become standardised through normative rules, the desire to adhere to these norms should not be afforded priority over the conduct of the research itself or achieving logical clarity in its depiction.

1. Introduction

Consider the following question: *When an individual or company sets out to design a product for addressing an identified security threat, what value is there in them looking beyond the seemingly overriding inquiry; 'does my product work'?* Before one posits an answer to this question, it is worth reflecting upon the following two events:

On the 14th of November 2011 the European Commission adopted new rules governing the use of body-scanners[1] within European Union airports that effectively banned the backscatter variant due to its employment of x-rays (European Union 2011). In the United Kingdom, Manchester Airport had invested heavily in implementing this technology by making it central to their passenger screening processes. But despite the costs incurred, by October 2012 they too had removed all backscatter scanners at their own expense (BBC News 2012).

Achieving the same effect, on the 18th of January 2013 with little prior public warning the United States' Transportation Security Administration announced they had cancelled Rapiscan's contract to provide backscatter body-scanners in US airports. As a result all deployed machines, 174 in total, were subsequently removed from thirty airports nationwide (Guardian 2013).

What makes these two events noteworthy are the justifications behind them. They are not based on a particular instance of catastrophic failure; for example where a passenger successfully smuggled concealed explosives through a backscatter scanner and subsequently detonated them mid-flight. Additionally, unlike previously removed technologies they are not justified by the failure of these scanners to reliably achieve their *raisons d'être*[2], and neither are they based on a wider societal rejection of all body-imaging technologies; while backscatter scanners were removed from airports, the alternative millimetre-wave body-scanner variant remains.

[1] Referred to as *security scanners* by the European Commission.

[2] Explosive trace portals (ETPs), for example, were systematically removed from US airports after failing to meet operational and reliability standards. Discrepancies between laboratory and operational testing-results, their inability to cope with dust and humidity in airports, and their propensity for regularly breaking-down all contributed to this scenario (see GAO 2009).

Ultimately, these decisions are borne out of the social controversies and resistance that arise in response to the introduction of these scanners; resistance which in turn is a response to the design choices made by Rapiscan, the developer of this technology. Specifically this entails; (a) *privacy concerns* based on the nature of the graphic images Rapiscan decided their scanners should produce[3], and (b) *health concerns* intrinsic to the fundamental decision to employ x-ray technology as the basis for their scanners.

This *failure* of backscatter body-scanners does not stand isolated; rather it is one of many examples where aspects of the design of a security technology either directly or indirectly evoke controversies within societies[4]. It is the design of controversial security technologies such as these, and my refusal to accept the occurrence of such social responses as a *fait accompli* that are the driving forces behind my dissertation. Given that the *shape* of technologies reflect the design choices of those who create them, rather than being predetermined, the fundamental question addressed within my dissertation is *how can we design socially acceptable security technologies?*

Before I commence a discussion of the aims, objectives, and methodologies of my PhD, it is necessary to define a number of terms employed throughout this dissertation. Specifically these are; *security technologies*, *society*, and *societal response*.

By *security technologies* (STs) I am referring to 'the product of an engineering endeavour that seeks to deter, prevent or detect crimes, and/or enhance the security of individuals, their property, or the state (including its infrastructure)'. This may include potentially lethal technologies, but for inclusion a technology must be available to civilian agencies and not restricted to the military. It is not restricted to technologies designed and/or developed by individuals holding formal qualifications, such as engineers, scientists, or industrial designers. Finally it encompasses all forms of *engineering pursuit* so long as what is produced has either a physical or digital presence. For example, while a bike-lock, body-scanner, or security centric data-mining programme are all examples of security technologies, a policy, law, or general computer operating system do not suffice.

[3] See Figure 2.1, page 64, for an example of this image.
[4] Chapter 2 of this dissertation examines twelve such examples, including whole body-scanners.

When I employ the term *society*, I define this to mean 'the situation of being in the company of, or associated with, other people through some determinable measure of proximity, or some shared trait(s) or belief(s), or external commonality'. In not restricting my definition to a large homogenous unit (i.e., British society) I also acknowledge that pluralities exist through the facts that; (a) an individual will simultaneously belong to many *societies* (e.g., UK citizen, male, Muslim, minor, white, etc.), and (b) different *societies* may react differently to, have different opinions of, and be treated differently by, the same ST.

By *societal response* I am referring to the future acceptability of a proposed ST to society. However, to keep my research project manageable I am limiting the societies I am contemplating to Western cultures; primarily the United Kingdom[5], though examples are included from the United States, Australia, and Northern European nations.

1.1 Research aim and objectives

The overarching aim of my research project is:

To assist the developers and designers of security technologies in anticipating and mitigating negative societal responses to their technologies upstream in the design process.

To achieve this overarching aim the following three primary objectives (with their accompanying requirements where appropriate) are as follows:

1. To identify the causes of social controversies arising from the design and operation of STs.

To consciously address a potential source of future controversy within the design of a ST before that technology is completed and deployed requires its developers explicitly identify these controversial design elements. This identification process would be both simplified and standardisable if individual controversies were not bespoke to each ST

[5] In the conduct of my research, and the subsequent production of this thesis, my focus of enquiries is primarily centred on the UK. My arguments and perspectives throughout should be read to reflect this fact.

but arose within diverse STs[6]. As such I endeavour to identify whether common controversies arise repeatedly across a range of diverse STs, thus allowing for their classification within an appropriate taxonomy.

2. To generate guidelines for the assessment/development of future design tools produced to assist those developing STs, such that these tools can adequately identify potential sources of social controversies within the unique field of ST design.

For future design-assisting tools to possess immediate relevance they must reflect the existing conditions prevalent within the security industry. This requires knowledge of the skills and training of those science, technology, engineering, and mathematics (STEM) practitioners currently designing STs, their common organisational work structures, and rules enforced by state organisations. Acquiring this knowledge requires interviewing individual engineers and scientists actively engaged in the design and development of these technologies. This information is analysed and transformed into general guidelines used to both assess the usefulness of existing methodological approaches and to guide the development of new tools[7].

3. To begin the process of creating these methodological design tools.

Finally I begin the process of applying the results from the first two objectives to the creation of methodological design tools for upstream-use by those scientists and engineers engaged in the process of designing and creating STs. Ultimately this requires the creation of entirely new design tools.

I would contend at this point that measuring the success of this research project is not determined by any one of these three objectives, and in particular the creation of a successful design tool. Arguably this output (constituting a tangible product applicable to *real world* ST design projects) would be the obvious yardstick for determining success. However, the value in identifying and classifying into a single taxonomy the

[6] For if no such commonalities exist, or if they do but are only relevant to specifiable categories of STs at a narrow level of abstraction, then future design tools will have to either be created differently for each individual category of ST or, in the most labour-intensive scenario, constitute a bespoke creation for each design project.

[7] The alternative (equally valid) approach would be to assess existing organisational and governance structures for weaknesses that contribute to the prevalence of social controversies arising from STs. I could then firstly recommend changes to these structures, then base any future tool design on this alternate reality. However, as discussed in Chapter 1.3 below I have chosen to ground this research project firmly within existing structures and constraints.

existence of common controversies that apply *across* the diverse spectrum of STs, as opposed to focussing on specific ST sub-groups[8] (such as ICTs, CCTV, surveillance technologies, less lethal weapons, etc.) in itself represents a novel, valuable contribution to the wider ST literature. Equally the insights gained from the process of investigating whether it is even possible to create such design tools is as equally valuable as the production of any final product. Additionally, the results obtained by interviewing STEM practitioners engaged in the process of developing STs provides insights into the working practices of a group rarely opened up to examination due to the impact of secrecy[9]. The creation of guidelines for the construction of design tools based on these interview results[10] possesses applicability beyond any tools created within my research project. In applying these guidelines, future design tools will also benefit from being grounded in existing corporate- and state-governance frameworks, as opposed to a counterfactual relying on changes to these frameworks that cannot be guaranteed[11].

1.1.1 Basic structure and methodology

The structure of my research project, as depicted in Figure 1.1, is divided into five distinct, sequential occurring stages, each relating to specific chapters within my dissertation. These are briefly outlined below. Please note that detailed discussion of each stage's methodologies and methods are included in each individual chapter, rather than being collated into a single methodology section here in the introduction, as a consequence of the particularly diverse formats of these stages.

The content of each stage is as follows:

Stage 1. *Case-study analysis of controversial security technologies* (see Chapter 2): Involves conducting *multiple document-based case-studies* of twelve selected STs that have all evoked social controversies, with the aim of identifying whether commonalities existed between these controversies regardless of the nature of the

[8] As achieved in Stage 1 of this research project, presented in Chapter 2.
[9] See Chapter 1.4 for a discussion of *secrecy* and Chapters 3 & 7 for interview results.
[10] See Chapter 6.1 for these identified guidelines.
[11] See Chapter 1.3 below.

individual STs themselves. The presence of commonalities that transcend the spectrum of STs is successfully identified. As a result of this process, forty-three *commonalities of controversy* are identified and subsequently organised into seven categories to produce a taxonomy of ST controversies[12]. More granular information pertaining to each of these forty-three commonalities is also derived from the case-studies, for use in future tool designs.

Stage 2. *Interviews with engineers and scientists* (see Chapter 3): Involves conducting qualitative, *semi-structured interviews* with fifteen STEM practitioners engaged in designing and developing STs, with the aim of understanding how this process occurs. This stage elicits valuable insights into both the motivations of these individuals, their skill-sets, and the constraints inherent to working within the ST industry.

Stage 3. *Case-studies of existing tools/methods* (see Chapters 4 & 5): In Chapter 4, the results of the case studies from Stage 1 and the interviews from Stage 2 are combined to produce a set of assessment criteria by which to judge methodological tools seeking to identify and mitigate negative social reactions to STs during their design process. Chapter 5 involves focussed case studies of a selection of existing models, approaches, and assessment tools. These are then compared against the assessment criteria produced in Chapter 4 with the aim of identifying suitable candidates for use as *off-the-shelf* design tools capable of being applied to all STs. Unfortunately, based on this process, no suitable candidates are identified. This necessitates the research undertaken in Stage 4.

Stage 4. *Creation of design rules and bespoke tools* (see Chapter 6): Through combining the work undertaken in the previous three stages, Stage 4 involves the production of design rules for governing the creation of future design tools. These are then applied to produce two methodological tools for assisting the designers of security technologies in anticipating and mitigating possible negative social reactions to their future products upstream in their design processes. These are:

- Framework of Common Controversies within Security Technologies (FCC)
- Designing for Socially Acceptable Security Technologies (DeSAST)

[12] See Table 2.13 Completed Taxonomy of Security Technology Controversies

Both of these, and their antecedent design rules, are presented in detail in Chapter 6.

Stage 5. *Assessment of created tools* (see Chapter 7): Involves a non-comprehensive evaluation and validation process of one of the design tools created in Stage 4 whereby DeSAST was applied to a ST currently under development by one of its designing engineers[13]. Observations from this application and an accompanying interview are included in Chapter 7, along with preliminary observations for improving DeSAST. This process also identifies another potential use for DeSAST; that being as an interview tool to facilitate sociological and ethical investigation into the thought processes, prioritisations, considerations, and justifications employed by STEM practitioners when contemplating the design of a future ST.

[13] That being the utilisation of Wi-Fi as a source of radar. See Chapter 7.2 for details.

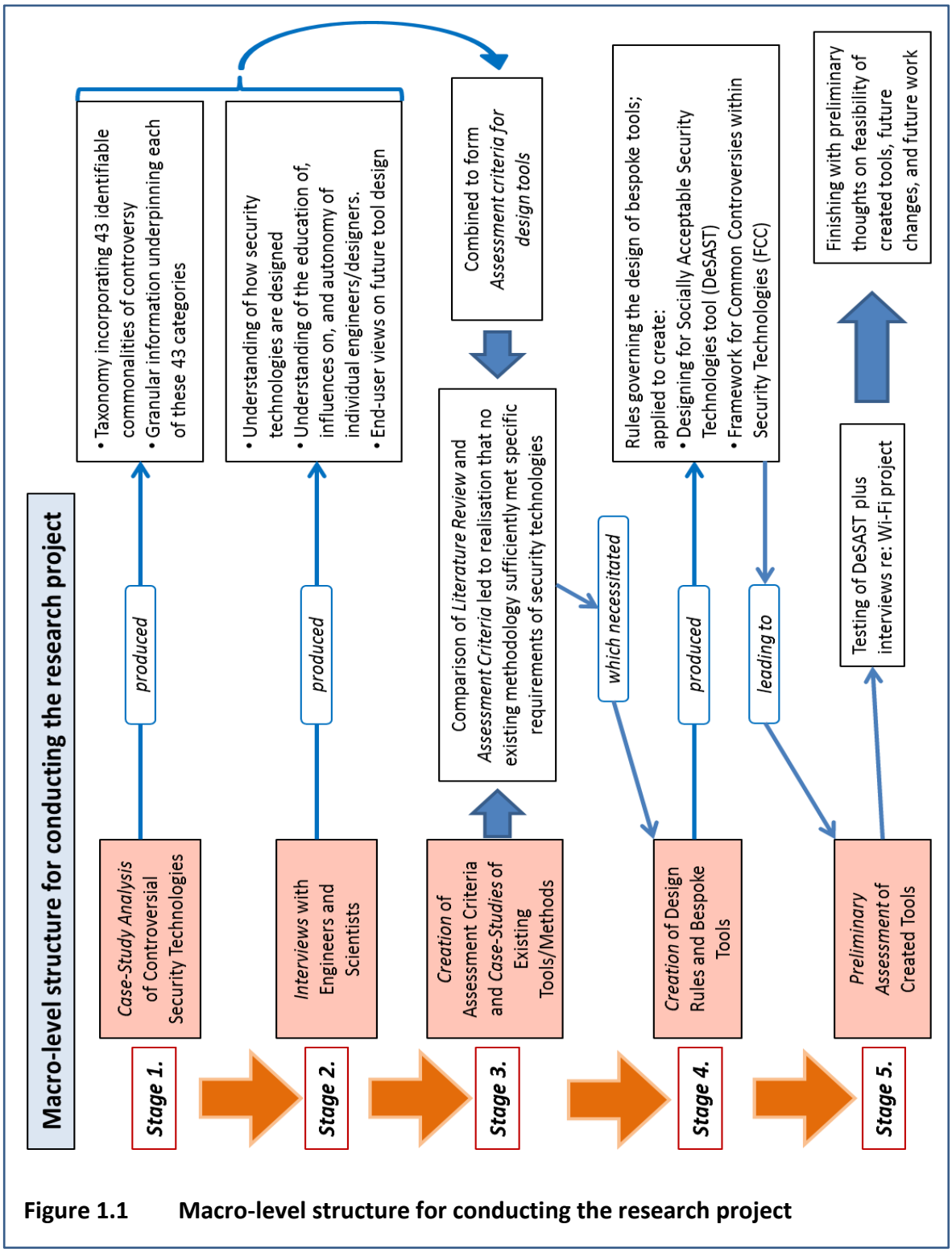


Figure 1.1 **Macro-level structure for conducting the research project**

1.2 Why this topic and why the narrow focus?

Before progressing further it is worthwhile addressing two questions at this point. Firstly *why focus on STs and in particular their social acceptability*? And secondly *why narrow this focus even further to only concern myself with the upstream role of STEM practitioners engaged in designing and developing these technologies?*

In response to the first question, the simple and concise answer is: 'because there is a need for focussed research on the social acceptability of STs'. To expand this response; I focuss on STs due to their increasing prevalence as instruments for deterring, preventing, detecting, investigating, and prosecuting criminal activity (see Roberts 2011; PERF 2012; den Boer 2011 and others for examples of this phenomenon). They are already essential tools employed by law enforcement and national security agencies, and yet their full potential has by no means been tapped. For as our collective body of scientific knowledge grows, as new technologies and manufacturing processes are developed and refined, and as what we can achieve on a digital level continues to expand, we will possess the ability to create previously unachievable STs. Furthermore, the terrorist attacks both on and after September 11th 2001 (9/11) have provided an impetus (by way of opportunities and rationales) for investing in, researching, developing, and implementing new STs on scales and at speeds that might not have been possible otherwise (Lyon 2006; Schneier 2006). These factors all serve to highlight the importance of maximising social acceptability in relation to these STs.

In relation to the second question, given the vast array of areas available and open to study it is sensible to narrow my focus. This decision to focus on both the upstream and the designer is justified for the following reasons. Firstly by failing to anticipate and address socially unacceptable elements within a new ST's design, these elements can undermine what may otherwise have been an acceptable ST[14]. In other words, attempting to address problems after deployment can prove too late. Secondly the final design of any ST is not preordained to assume a certain form or achieve its functional goals in a certain way[15]; rather it is the result of design choices by the

[14] See Chapter 2.2.2 for a discussion of the failure of backscatter body scanners in US airports because of privacy concerns arising from initial design decisions about the images produced, and the developer's subsequent inability to address these issues.

[15] See Chapter 1.5.1 for a discussion on the rejection of technological determinism.

designers. As such the upstream developers of these technologies both possess the scientific/technical skills and inhabit a privileged physical location to shape the form of those STs they are creating. However, it does not follow that they correspondingly possess the necessary skills, focus, tools, or training to identify potential future sources of social controversy within their designs, nor that their workplace organisational structures enable such actions[16]. As such, focussing on this specific group of actors, with the aim of assisting them in identifying potential sources of social resistance within the future products they are tasked with creating, becomes a logical, justifiable approach even if it only tackles one link in a much larger chain.

One final reason for opting to conduct research focussed on the social controversies arising from STs, without limiting the scope of this enquiry to specified categories or specific examples of a ST[17] or to specific types of controversies[18], is the scarcity of existing literature adopting this approach. Existing approaches have tended to restrict the focus of their enquiries by specific ST, specific controversy, or a combination the two. While this approach is valuable, I believe there is increasingly a place here for *also conducting* more holistic examinations. Individual STs are increasingly being linked together into single platforms (e.g. closed circuit television (CCTV) combining with facial-recognition systems, behaviour algorithms, automated threat detection systems, automated number plate recognition, persistent authentications systems, etc.) and/or together into larger networks of multiple STs. As such, approaches designed to assist ST designers that are limited to a specific technology may find decreasing applicability here. Similarly the diverse nature of social controversies arising from STs makes it risky for the developer of a ST to limit what they seek to correct for in their designs, especially given the potential consequences of failing to address the source of a future social controversy[19].

[16] See Chapters 3.3 & 3.4 for a discussion of the results from interviews conducted with STEM practitioners engaged in the process of designing STs whereby these skills and organisational limitations are identified.

[17] For example; CCTV (Norris and Armstrong 1999), data mining and profiling technologies (Custers et al 2012), less lethal weapons (Rappert 2003), etc.

[18] For example; privacy issues (see Solove 2008), data protection issues (see Akrivopoulou and Psygkas 2011), human rights issues (see Cunha et al 2013), etc.

[19] See Chapter 6.1 for an expanded discussion justifying this holistic position.

Isolating and focussing on the *upstream* role of those engaged in the development of future STs (predominantly STEM practitioners) does not deny the importance of *downstream* actors/actions in affecting social acceptability. Insupportable rules of governance by elected officials, overuse, unacceptable use, and/or abuse by end-users all possess the propensity to undermine the public's perception of otherwise acceptable STs[20]. But the fact is there is no *silver-bullet approach* to ensuring the social acceptability of any ST, and as such all aspects of the lifecycles of these technologies are necessary targets for research.

1.3 A conscious decision to work within the world as it is

The decision to focus my research on assisting the developers and designers of security technologies to anticipate and mitigate negative societal responses to their technologies upstream in the design process, necessarily generates the need to address another fundamental issue. That being: whether to ground any assistance measures in the existing *structures* [21] governing both the design of STs and the operation of the security industry, or whether to first produce new idealised structures and then base the design of any assistance measures on these.

I have chosen to ground my research project in the existing structures governing the security industry[22], characterised as the following:

- The domination of secrecy throughout, actively enforced by governments. A system which pervades the design and development of STs[23]. Arguments often

[20] An example being the publically perceived misuse/overuse of surveillance powers under the Regulation of Investigatory Powers Act 2000 (RIPA) whereby covert surveillance technologies were employed by various Councils to police school catchment zones, dog-fouling, and littering (BBC News 2008; BBC News 2010).

[21] *Structures* here refer to the various rules (formal and informal) which determine how STs are currently designed and developed within a country. This includes all legal provisions, codes of practice, and guidance provided by governments, as well as the contractual, organisational, and normative rules operating within the ST R&D industry as a whole.

[22] In doing so I am focussing here on the UK. It must also be noted here that the concept of 'the security industry' is "*neither well defined nor clearly identifiable*" (ECORYS 2009, p.i). Blurring is identified between *traditional* external security (characterised by defence) and *new* internal security characterised by counterterrorism and law enforcement. External conflicts and global organised crime effect internal security, internal security threats may require military support, and STs may possess both internal/external dual-use applicability (ECORYS 2009). For my purposes when referring to 'the security industry' I am only including those segments of the industry whose products are not solely for military use.

used to justify this secrecy include; national security, commercial sensitive information is involved, projects may involve partners in more than one country, revealing information will compromise the effectiveness of technologies and procedures[24].

- The virtual absence of public engagement prior to the initial introduction or announcement of a ST[25].
- A commercial industry driven almost entirely by profit rather than any sense of altruism or national-duty[26].
- A competitive industry dominated by large operators, with small and medium companies typically restricted to niche markets or the licensing of their products to larger operators (ECORYS 2009).
- The predominance of STEM practitioners as designers and developers of new STs.

In so doing *I am not endorsing these characteristics*, especially the secrecy requirements adding to the lack of public engagement. Nor am I claiming such structures are optimal for minimising the propensity of STs to result in social controversies. Indeed the entire focus of my research project represents a tacit acknowledgement that there are serious flaws embedded in these existing 'closed-shop' design processes. However, by working under the premises that; (a) the existing structures will probably remain largely unchanged in the short to mid-term, and (b) if they do change I have no way of accurately predicting their new forms, I am maximising the contemporary relevance of this research project by accepting the ST world in its current form.

[23] See Chapter 1.4 for further discussion of this phenomenon.

[24] See Wright and Raab (2012) for a critical examination of these and other arguments within the context of surveillance activities.

[25] See Chapter 1.4 for further discussions. Notable exceptions do exist here, such as the public relations campaign waged by the UK government when proposing the introduction of identity cards under the Identity Cards Act 2006 (ICA).

[26] This ostensibly excludes government departments/agencies utilising solely in-house employees and expertise to produce STs, but even this becomes blurred when governments seek to license or sell products at a later date.

1.4 Secrecy within security technologies

If I were ever required to identify *the one single concept* with the greatest impact on my dissertation it would be *secrecy*. In a variety of forms, secrecy repeatedly arises within, and often dictates the form of, my entire PhD; be this in the design of STs, the conduct of my data collection (especially the anonymity requirements insisted upon by the STEM practitioners interviewed in Chapter 3), or the construction of the design tools for assisting ST designers/developers undertaken in Chapter 6. As such, what follows in Chapter 1.4 is a detailed discussion of this concept, its impact on my research topic, and how it has been accommodated for throughout my dissertation.

Secrecy, defined by Bok (1989) as *deliberate concealment*, is a commonplace concept within the realms of national security and law enforcement, encompassing such diverse topics as the conduct of undercover operations, covert surveillance, the membership of various government agencies, and the use of *closed material procedures*[27] amongst others. Secrecy of scientific and technological knowledge possessing national security and/or military implications is also an established ideal (see Rappert (2007) and Balmer (2012) as examples of different contextual accounts).

When it comes to STs, secrecy pervades their design, development, deployment, and continued operation. However, this claim embodies all manner of contradictions. Firstly private companies producing their own STs for commercial sale will generally be able to decide how open they will be regarding both the design process and information pertaining to their products. But, if they are working under contract and have signed secrecy requirements then their freedom to divulge information will be constrained. Secondly secrecy is neither universally applied nor required within STs, rather it operates on a spectrum as depicted below:

[27] Commonly referred to as 'secret courts' these are procedures under the UK's Justice and Security Act 2013 whereby evidence in civil cases is not disclosed to claimants on national security grounds (Bowcott 2013).

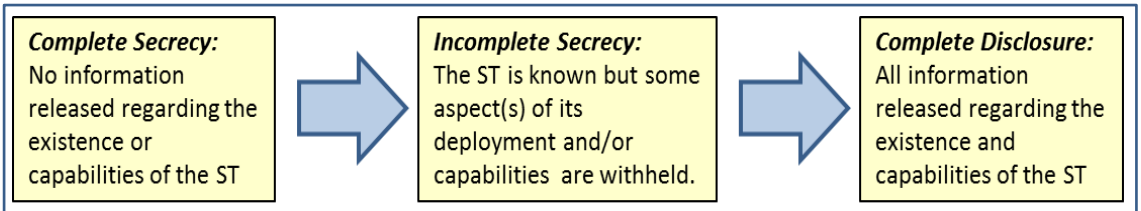


Figure 1.2 Secrecy within STs

To provide examples, at one end of this spectrum international mass information and communications technology (ICT) surveillance programmes such as PRISM[28], have been designed, developed, implemented, and operated under a blanket of *complete secrecy*. The *incomplete secrecy* category includes whole-body scanners whose existence and purpose are disclosed but whose detection capabilities/rates remain classified. Finally there are STs where *complete disclosure* is employed, such as simple bike locks and chains whose resistance to being cut is made public.

These different approaches feed into a set of contradictions centred on the purpose(s) of any ST. Is the ST employed for deterring, preventing, detecting, investigating, and/or prosecuting criminal activities? Depending on what it is the end-user is hoping to achieve with a ST, the presence of secrecy can aid or hinder these goals. Using the examples above, data mining programmes like PRISM can be used to prevent, detect, and investigate criminal acts, but while they remain completely secret they cannot achieve direct deterrence. However, in regards to whole-body scanners, by implementing these in airports end-users can use them to prevent, detect, as well as *deter* and prosecute criminal activities. That said, it is arguable that the deterrence effect is maintained by *both* the scanner's presence *and* the fact that information on just how effective these devices are at detecting non-metallic and organic materials is not released to potential attackers – thus making it harder for them to circumvent this technology[29]. Finally the manufacturer of a bike lock/chain for preventing bike theft may seek to be completely open about their product's capabilities, thus relying entirely on deterrence and prevention, even though the thief has now been provided all the

[28] See Greenwald and MacAskill (2013).

[29] There are numerous caveats to this claim. Firstly if in reality the classified detection capabilities of this ST are in fact very poor then its effectiveness relies entirely on its deterrence effect; what Bruce Schneier (2006) refers to as *security theatre*. Secondly, where the deterrence effect of a ST relies upon secrecy as to its capabilities, should attackers gain access to this information and circumvent this ST without the knowledge of the defenders then that ST will no longer be an effective deterrent.

information they need to successfully carry out an attack (i.e., the equipment and time required to break the lock).

Despite the ostensibly legitimate applications of secrecy within STs described in the above examples, secrecy can also be employed to further rather more questionable ends. As noted by Balmer (2012), secrecy can act to prevent the public disclosure of information with the potential to result in public censure or exacerbate existing public resistance. This discussion on the application of secrecy goes beyond ST design, reflecting the wider debate of how to reconcile *secrecy* with *democracy*, where these two concepts are often presented as conflicting ideals. However, such a simplistic opposing conceptualisation masks a more nuanced debate. As Thompson (1999) points out, "*[s]ome of the best reasons for secrecy rest on the very same democratic values that argue against secrecy…. The conflict involves this basic dilemma of accountability: democracy requires publicity, but some democratic policies require secrecy*" (p.182).

Thompson (1999) accepts as justifiable the need for government secrecy but with three important caveats. Firstly there should be a presumption of publicity/transparency, such that in any balancing of publicity and secrecy, publicity should be afforded primacy. Secondly, that if a reason a policy is kept secret is because it would be defeated by democratic processes if it was made public then that policy should be abandoned. Thirdly, for secrecy to be justifiable, the process by which this decision is justified must not itself be a secret one. Thompson advocates as democratically acceptable a concept of *obscurity*: situated "*somewhere between deep concealment and full disclosure. Such secrets are not completely concealed …. But their content is not made explicit … [which is] necessary for the policy's being effective*" (p.186), which corresponds with the concept of *incomplete security* within Figure 1.2 above.

1.4.1 Secrecy and conducting research

At the level of the STEM practitioner, secrecy can pose fundamental challenges to successfully undertaking research. By its very nature it prevents a research interacting with other colleagues not afforded access to the project at hand, which in itself can

slow the rate of scientific progress purely by restricting sources of potential input when seeking to solve problems. That said, Rappert and Balmer (2007) note that within military research establishments epitomised by secrecy (such as those engaged in the development of weapons of mass destruction), shadow communities of scientists have been encouraged to form which, amongst other things, assist information flows. Despite the possibility of such networks forming, the development of STs within private companies may pose additional challenges if the prioritisation of commercial secrecy is used as a justification for denying STEM employees access to even these shadow networks[30].

Another consequence of enforced secrecy is the isolation of scientists, not only from peers, but from those individuals and groups (including fellow scientists) morally opposed to the work being undertaken (Balmer 2012). This can result in a form of moral and ethical bounded rationality, whereby the negative aspects/consequences of one's work are either diminished in importance or are simply not recognised[31].

In addition there are the (potentially severe) personal implications for a ST designer who subsequently rejects secrecy by violating national secrecy provisions. The US Government's vigorous pursuit of computer programmer and former National Security Agency 'contractor turned whistle-blower' Edward Snowden, for the "*unauthorized communication of national defense information*" and "*willful communication of classified intelligence with an unauthorized person*" (Finn and Horwitz 2013) under the US Espionage Act 1917[32] provides a poignant case-in-point.

As a concluding point to this section, despite the negative aspects of the above discussion it must be noted that within the specific constraints of this research project, *secrecy* may actually assist in the operation of my proposed design-tools. If the intended end-users of these tools apply them knowing their responses will be protected under a cover of secrecy, then they may approach their task with greater

[30] To some extent, staff changing employers and moving between ST companies will involve a level of knowledge transfer, though its speed and scope will be limited.

[31] For an example of this phenomenon occurring within this research project see Chapter 7.2.1 under the subheading: *Considering a rejected section*.

[32] As incorporated into the United States Code; and more specifically Title 18, Part 1, Chapter 37, Section 798.

candour, confident that their actions and responses will not be divulged to unauthorised outsiders.

1.4.2 Direct implications for this project

The effects of secrecy manifest themselves primarily within two elements of this research project: (1) the interview components, and (2) the composition of the design requirements I produce to governing future tool design.

Regarding my interviews with the fifteen STEM practitioners conducted in Stage 2 (and discussed in detail in Chapter 3.2.1), due to the restrictions placed on these individuals under legislation such as the UK's Official Secrets Act 1989 and the consequences of breaching this Act, my interviewees were understandably cautious about both agreeing to be interviewed and the reporting of their comments. As a result they stipulated (and have been provided) full anonymity throughout my dissertation, including the complete disassociation of comments/quotes from the individual sources. These stipulations for anonymity also impact the format of the interviewee consent form. As presented in Appendix *A*, this form is somewhat unconventional (though not without precedent[33]) in that it is not intended or designed to be signed by the interviewees, an action that would have undermined anonymity. Rather its primary purpose is to expressly set down how I (the researcher) will protect the interviewees when collecting, storing, and presenting their interview data. I base the decision to proceed with completely anonymised data collection in the absence of signed consent forms on the following:

1. Consent is still obtained through the particular consent forms I do use[34] which clearly set out the protections afforded the interviewees.
2. The potential benefits of conducting this research (in the form of more responsibly designed STs) outweigh the costs of keeping total anonymity.
3. Collecting and analysing anonymised data, stripped of identifying features, is better than not being able to collect any data at all.

[33] See Farrimond (2013) for a good discussion of this problem.
[34] See Appendix *A*

Regarding the design requirements presented in Chapter 6.1, the effect of secrecy translated into Requirement 4: *The tool must be usable without any direct public interaction or input*. Thus rather than challenging the efficacy or legitimacy of excluding the public from ST design projects I have chosen to create a requirement that any future design tool must be usable without such input. By including this requirement I am not making the obviously false claim that *all* STs are developed under a secrecy requirement. Nor does this mean I am advocating the creation of design tools which expressly prohibit the designers/developers of a STs from engaging with the public should they have permission to do so. Rather this requirement simply seeks to ensure that any future design tool is able to operate effectively in an environment where public engagement is not possible/permitted.

1.5 Factors pertaining to, and influencing, the theoretical positioning of my research

1.5.1 A rejection of technological determinism

Martin and Schinzinger (1996) highlight the potential for theories developed within Science and Technology Studies providing important insights into engineering ethics and practice; in particular the rejection of technological determinism. Technological determinism is the view that technology develops independently from society via some internal logic, and upon its release it affects the character of society (MacKenzie and Wajcman 1999). Science and Technology Studies orthodoxy holds that there is influence and shaping passing *both ways* between technology and society, beyond simple cause-and-effect mechanisms, which affect the development and nature of both. On a practical level, if this was not the case (such that the development trajectory of a ST was entirely predetermined by technical factors and incapable of social influence) then this research project would have no justifiable theoretical basis.

However, this inclusion of social factors does not negate the *real* restrictions inherent within science and engineering. While Science and Technology Studies seeks to draw out into the open the important social components of technoscience, in a discussion

on engineering Vincenti (1995) warns his fellow practitioners of privileging the social over the technical (or vice versa) when examining the activities of engineers. Vincenti makes clear that one can accept the importance of social construction to the shaping of technology without abandoning the fact a real world exists beyond human wishes; a world that that "*imposes intractable, non-negotiable constraints on what engineers can and cannot do*" (p.553).

The important implications of social shaping on my research are that engineers are not merely unearthing a predetermined technological artefact, devoid of social influences, assumptions, and biases, when creating a ST. That STs do not simply shape a society upon their deployment, rather they are also being interpreted, constructed, and shaped by that same society. And finally Science and Technology Studies perspectives imply that because the designers of a ST have discretion in what they produce they cannot escape broad ethical responsibilities for their work.

1.5.2 Persevering with STs as technological fixes to crime[35]

The development and interpretation of new technological advancements are incorporated into STs with considerable enthusiasm by governments, law enforcement agencies, and private companies as potential methods for preventing, detecting, and prosecuting criminal activities. In this regard STs represent *technological fixes* for the social problem of crime; a technological fix being broadly defined as a technological solution for *solving* social problems (Weinberg 1967).

STs are often presented as a panacea for addressing crime by being cheaper and/or more effective than the alternative human-centric approaches. Whole body scanners at airports utilise X-ray backscattering or millimetre wave technology so as to identify metallic *and* non-metallic objects, plastic and liquid explosives, flora, fauna, drugs, and cash, concealed within or beneath the clothing of passengers (European Commission, 2010; Mitchener-Nissen et al 2012). Data mining, being the application of database

[35] Barring some modification I have published the text of Chapter 1.5.2 in the following article during the course of my research project: Mitchener-Nissen, T (2013) Addressing social resistance in emerging security technologies. *Frontiers in Human Neuroscience*, August 2013 (7), doi: 10.3389/fnhum.2013.00483

technology and techniques (such as modelling and statistical analysis) to data so as to identify valid, novel, implicit and potentially useful information and patterns within that data, is employed with the aim of analysing intelligence and detecting terrorist activities, fraud, and other criminal patterns (Schermer 2011; Steinbock 2005; Tien 2004). The use of biometrics enables crime-scene technologies that can assist in the identification and prosecution of offenders (such as DNA databases and fingerprinting technologies), tackling identity fraud, and counteracting illegal immigration (Goldstein et al 2008; Grijpink 2006). To assist in the investigation and prosecution of criminal acts, researchers and private companies are trying to develop lie detection technologies designed to directly access brain functions by employing fMRI and EEG (Wolpe et al 2010). This selection represents a tiny snapshot of the cornucopia of STs both under development and already implemented.

Without further examination it is tempting to conclude that STs do indeed constitute a justification for Weinberg's vision of technological fixes as the solution to social problems. However the notion of the technological fix is subject to robust criticism. It is also described as "*a quick cheap fix using inappropriate technology that creates more problems than it solves*" (Rosner 2004). The truth of this statement is evident within the social controversies (or in the case of the lie detection technologies, the possible future social controversies) produced by each of the ST examples provided above. Whole body scanners have been accused of conducting digital strip-searches (Klitou 2008), and the backscatter variation has been removed from US airports because of the images produced. Data mining has been associated with both a fear of totalitarian-style state observation, as well as the targeting of individuals by governments (Steinbock 2005). Different biometric technologies can discriminate against various groups within society and are plagued by the problem of false positives (Hunter 2005; Whitley and Hosein 2010). Additionally the UK's DNA database (the largest in the world) has created controversy by holding the details of innocent people and a disproportionate number of samples from ethnic minorities (Jobling and Gill 2004; Human Genetics Commission 2009). And finally the new generation of potential lie-detection technologies have faced criticism over the potential ethical, social, and legal implications of their operation to existing social and legal institutions should they

ever be made to definitively and consistently 'work' (Rusconi and Mitchener-Nissen 2013).

Despite the *new problems* created by these technological fixes (i.e., by STs), I have not discerned any reduction in governments' appetites for their development and deployment. On the contrary, rather than questions being raised over the wisdom of persevering with the 'ST as technological fix' approach, quite the opposite response is observable. To fix the social problems created by the technological fixes that are STs, what are often developed are *additional* technological fixes; i.e., *new technological fixes* to fix the *new social problems* created by the *old technological fix* created to address the *old social problem* of crime. For example; developing privacy enhancing technologies (PETs) such as image modifiers[36] to address the privacy concerns[37] created by the whole-body scanners[38] that were introduced to address the problem of people smuggling non-metallic items concealed within or under their clothing onto aircraft[39]. The definitive message to be taken away here is that we should not expect the creation and deployment of STs to diminish in the near future given the enthusiasm of their adoption, regardless of the associated problems they may create.

1.5.3 Imposing additional ethical responsibilities on engineers/scientists

By focussing on the upstream development of STs, while maintaining the requirement of secrecy, a concomitant implication is the imposition of additional responsibilities onto the STEM practitioners involved in creating these technologies. These individuals will be held accountable for identifying and addressing sources of social controversy within their products. This approach raises two related questions:

1. Is it appropriate to impose this ethical burden on these STEM practitioners?
2. How equipped are they for achieving this task?

Regarding question 1, the referenced *ethical burden* essentially requires the designers of a ST accept responsibility for the social controversies their products subsequently

[36] Representing a *new technological fix*.
[37] Representing the *new social problem*.
[38] Representing the *old technological fix*.
[39] Representing the *old social problem*.

evoke. This is admittedly a contentious requirement, though in relation to engineers it reflects modern trends and the culmination of gradual yet profound changes to their responsibilities over the past century. These have influenced how engineers conduct their business, particularly in relation to the development of engineering ethics, and the responsibilities they bear for the misuse or negative effects of their creations.

Examining the historical development of ethics in engineering from the formation of the Society of Civil Engineers in 1771 onwards, Mitcham (1997) identifies the emergence of three distinct ideas that influenced the conduct of engineers. The first idea is the primary duty of engineers to be *loyal to their employer*; an ideal formalised in the early ethics codes of the American Institute of Electrical Engineers in 1912 and the American Society of Civil Engineers in 1914. However, such a duty of obedience gives rise to the concern that it could open the engineer to unjust manipulation by an employer. Later, during the first third of the twentieth century, a second idea known as the *technocracy movement* arises. This holds that engineers should have political and economic power as they pursue technological efficiency, thereby allowing them to apply their own standards of good/bad and right/wrong without being beholden to business interests. The belief being this would result in better products for the consumer and a stronger economy (Mitcham 1997).

Yet pursuing technical perfection as a goal in and of itself has limitations and can result in human, societal, environmental, and other factors being ignored. Questioning of, and opposition to, technocracy began to grow post World War II giving rise to calls for *social responsibility* (Mitcham's third idea) as a component of ethics in engineering (Mitcham 1997).

Following a number of technological developments, designs and failures that negatively impacted human wellbeing throughout the period beginning with the development of atomic weapons[40], the academic field of *engineering ethics* was born in the early 1980s to assist engineers in making the right decisions when facing ethically difficult situations (Johnson and Wetmore 2008). This modern engineering ethics has been defined as:

[40] A sample includes; the development of nuclear weapons, nuclear reactor meltdowns, the widespread use of chlorofluorocarbons, DC-10 engines separating from the aircraft, the Union Carbide chemical plant disaster at Bhopal, and the dangerous fuel system design in the Ford Pinto.

> (1) the study of moral issues and decisions confronting individuals and organizations engaged in engineering and (2) the study of related questions about the moral ideals, character, policies and relationships of people and corporations involved in technological study (Martin and Schinzinger 1996, pp.2-3).

Martin and Schinzinger highlight the potential for theories developed with Science and Technology Studies to provide important insights into engineering ethics; the first being the rejection of technological determinism as discussed in Chapter 1.5.1 above. The second theory is the acceptance that engineers are not isolated individuals; rather they work within, and are influenced by, a network of actors that exert influence on the development of technological artefacts before, during and after the engineer ceases work (see Latour 1987 & 1996). As such the responsibility of the engineer for their creations, while important, is also diluted and redistributed (Johnson and Wetmore, 2008).

This leads into the vexed question of whether engineers should be responsible for the negative side-effects or misuse of their creations. Here there exist a wide range of responses, from the extreme view that engineers should be responsible for all the negative consequences of the devices they design and develop, through to the equally extreme opposing view that engineers should bear no responsibility for the misuse or negative effects of their inventions. This is another area in which Science and Technology Studies academics have been active.

Modern Science and Technology Studies and engineering ethics literature appears to adopt a moderate middle-ground. For example Whitbeck in *Ethics in Engineering Practice and Research* contends that while engineers do have a responsibility to make technology safe, this duty is restricted to foreseeable failures and misuses, for:

> [t]o be morally responsible for outcomes people must have some ability to foresee and influence them. I draw attention to this seemingly obvious point because some commentators have sought to blame technology (and engineers and scientists) for everything that is objectionable in modern life (1998, p.117).

Doorn and Fahlquist (2010) also list foreseeability as an accepted component of responsibility, before going on to raise the common problem of how to apportion responsibility when engineers work as part of a team whereby responsibility is diluted. They suggest embedding engineering ethicists within design teams at earlier stages, shifting the questions from ‘whom to blame’ or ‘how to apportion collective blame’ to

'how can we carry out the research such that ethics acts as a guide'. This approach can be contrasted to that of Johnson and Wetmore (2008), who contend that as engineers are engaged in building sociotechnical systems they still possess personal responsibility, but this responsibility shifts to become 'a need to communicate and coordinate with the other social actors'.

Despite these differences in determining the appropriate *nature* of the STEM practitioner's ethical burden, existing orthodoxy appears to display no compunction in *assigning* ethical burdens to these STEM practitioners in relation to their products. As a direct result I feel I am not unjustified in imposing a responsibility to identify and mitigate societal concerns.

While Question 1 at the start of this section sought to determine the *appropriateness* of imposing the ethical burden of identifying and mitigating societal concerns arising from the design elements of their creations onto STEM practitioners, Question 2 inquired as to how *well-equipped* STEM practitioners are for undertaking this task. On the *mitigation* requirement, I argue they are often the only people with the necessary skills to do so. Given the complex technical nature of many (but not all) STs, it requires trained engineers, natural scientists, computer programmers, and mathematicians to make the necessary changes and/or advancements to a ST's design so as to mitigate potential societal concerns. However, on the *identification* requirement I argue their ability to achieve this is severely hampered for a number of reasons[41].

The first is the paucity of social and ethical education within university STEM courses. From the interviews with STEM practitioners undertaken in Stage 2 of this research project[42] it was repeatedly highlighted that in university engineering, mathematics, and hard sciences courses in the UK, it is highly likely that a student can (and will) complete their education without ever undertaking a single lecture on the importance of identifying and incorporating social and ethics factors into their work. This is despite the creation of the field of engineering ethics as discussed above. For those who counter with the claim that ethics and ethical research is ensured by the presence of

[41] I have published the text of the first two reasons included here in the following article as part of my research project: Mitchener-Nissen, T. (2013) Addressing social resistance in emerging security technologies. *Frontiers in Human Neuroscience*, August 2013 (7), doi: 10.3389/fnhum.2013.00483

[42] See Chapter 3.

university ethics boards, while a particular research or design project may meet all official conduct requirements such that it is considered ethical, this does not mean that what is being undertaken or created will be accepted by the public. The diverse groups which comprise a society ultimately determine what is considered socially or ethically acceptable, and yet university engineering and hard science courses regularly fail new researchers and designers by not equipping them with an understanding of this fact nor the tools to adequately interact with the public.

The second element in the lack of priority afforded social and ethical issues within research and design projects. Again the interviews in Chapter 3 highlighted a clear hierarchical structure to the design process. For commercial projects it begins with cost; if it is determined that there is not a viable market for a product then it will not be produced. If this test is passed and the project is considered feasible than design specifications are produced in accordance with the client's requirements and the product is created. Similarly with university research projects, the presence of funding and/or the potential for future commercial exploitation dictates the research undertaken. When this is directed towards addressing perceived security deficiencies the focus is on attaining a specific security goal. These processes leave little space for the consideration and incorporation of social and ethical issues – the focus is on 'can we achieve what we have set out to achieve', and not 'is this a socially acceptable way of achieving the desired goals' or 'are these goals socially acceptable *per se'*.

The third is the combined effects of secrecy discussed in Chapter 1.4, including the restrictions on public participation within the design and development of STs, and the effects on scientists not being exposed to dissenting views challenging the underlying ethical values of their research projects.

1.5.4 The distinction between *legality* and *legitimacy* of STs

The fact that either democratically elected officials, or state agents/agencies who enjoy widespread popular support, that commission and/or support the development and deployment of ST, this does not prevent public resistance arising in relation to

these technologies. It is this fact that I contend highlights a fundamental division internal to STs; the contrast between *legality* and *legitimacy*.

The *legality* of any ST is determined by that technology meeting all the legal requirements (be they local, national, or international) for its use within a particular jurisdiction, as will be set by the political and legal structures governing that jurisdiction. In contrast, a ST attains *legitimacy* by being accepted by society at large. Legality cannot ensure legitimacy while legitimacy does not provide legality. But while legality can be created *before* a ST is designed deployed or made public, legitimacy by its nature can only be obtained *after* the public have been made aware of a ST, which requires the technology (or information on it) entering the public domain.

While legality and legitimacy are distinct concepts, the need for a ST to possess both cannot be overstated. For while possessing legality enables the introduction of a ST *into the field*, if that same ST does not also gain and maintain legitimacy through active public support or passive public acceptance *it will not survive unscathed* over the mid- to long-term. This definitive assertion is supported by the examinations undertaken within this research project of multiple STs that have all evoked social controversy that undermined their legitimacy[43]. As a result they have all been either prohibited, forcibly modified, their distribution restricted, and/or their usage constrained by various rules. The concomitant danger here is that by introducing these *socially unacceptable* technologies, trust in government and state agencies is threatened, research and design capacity is diverted from *acceptable* technologies, and money is wasted that could otherwise have been used for legitimate programmes. It has also been noted that the rejection of a technology can lead to its permanent inferiority through neglect (MacKenzie and Wajcman 1999).

As a result of these factors, once a ST that evokes social controversy is deployed, the negative implications arising from its lack of legitimacy will begin to take effect. The challenge, therefore, is identifying what aspects of a ST's design are more likely to evoke social controversy *before* a technology is developed and deployed. Once identified, the designers of a ST can consciously address these issues upstream in the design process. Thus they can attempt to mitigate future negative social reactions,

[43] See Chapter 2.

before these responses materialise, via explicit design decisions during the STs development process. This practical challenge forms the basis of this research project.

1.6 Precedents to the upstream approach

The upstream focus of this thesis, whereby attempts are made to influence the research and design of a technology by incorporating certain requirements and/or values, is not without precedent. From the *technology design perspective*, approaches such as *Value Sensitive Design* and *Privacy by Design* emerged during the 1990's and sought to infuse design processes with human-centric values and privacy protections respectively[44].

From the *technology governance perspective*, Article 23, Paragraph 2, of the proposed General Data Protection Regulation of the EU[45] adopts *as default* mechanisms for the processing of personal data which minimise the collection, processing, and retention of such data. These mechanisms (themselves comparable to STs) must also ensure that default settings prohibit the accessibility of personal data to an indefinite number of individuals. Adopting such a position pushes data protections upstream from individual users.

1.7 Adopting a mid-range approach to research

In an effort to both foster debate and challenge a perceived *status quo*, it was provocatively noted by Wyatt and Balmer (2007) that within Science and Technology Studies much of the work undertaken has tended to constitute either grand 'macro-level' theories or focussed 'micro-level' case studies. What was presented as lacking was a robust corpus of *mid-range* research; i.e., research that seeks to bridge the gap

[44] These two concepts are discussed in detail in Chapter 5.2.4

[45] EU Regulation on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) COM(2012) 11 final, which seeks to replace, consolidate, and strengthen the current EU data protections established under Directive 95/46/EC.

between grand theories and focused studies. By drawing on Robert Merton's[46] claims in respect of sociology, Wyatt and Balmer (2007, p.621) note that:

> [F]or Merton the failure to develop such theories, by focussing instead on the production of descriptions or the production of theories of everything, meant that sociology was unable to engage wider audiences with its work. For Merton, middle-range theory meant engaging with reality, albeit a limited aspect of it; producing theoretical accounts that engaged with that reality [that] … could be used to communicate with others.

Geels (2007) characterises a mid-range theory as one which "*[does] not address the whole of science and technology, but focus[es] on limited themes and topics. [One which makes] … explicit efforts to combine different concepts in an analytical model and … search[es] for patterns and explanatory mechanisms*" (p.635). While my dissertation is foremost a Security and Crime Science dissertation and not a Science and Technology Studies one, in my efforts to produce a multidisciplinary piece of research that would be accessible and relevant to a range of actors, I have adopted just such a mid-range approach.

Applying Geels' characterisation of a mid-range theory to my own research project; to move up from the micro-level I begin in Chapter 2 by exploring controversies within STs through repeatedly undertaking case studies of twelve controversial ST. This is complemented in Chapter 3 with individual interviews with fifteen STEM practitioners working in the field of ST design and development. From this data I seek to elevate the discussion beyond the *individual* (be that a technology or person) to a mid-range level with greater applicability. This is characterised by the identification of forty-three commonalities of controversy[47] and five essential design rules[48] that are subsequently combined in Chapter 6 to produce the two design tools (FCC and DeSAST). The analytical model produced in Chapter 2[49] represents an explicit effort to combine and incorporate different concepts (in the form of the commonalities of controversy) based on perceived patterns and a search for explanatory mechanisms.

[46] Specifically: Merton, R (1968) *Social theory and social structures*. New York: Macmillan.
[47] See Table 2.13
[48] See Chapter 6.1
[49] The Completed Taxonomy of Security Technology Controversies – Table 2.13

Similarly, in drawing down to the mid-level from the macro-level throughout my dissertation, I am not seeking to address the socially acceptable development of *all technologies*. Rather I am restricting my focus solely to a sub-group within the construction of technologies that is the social acceptable designing of STs.

2. Stage 1 – Case-Study Analysis of Controversial Security Technologies

The methodology adopted for Stage 1 entails undertaking individual case studies for 12 controversial STs (comprising Chapters 2.2-2.13 below), with the results of this process coded by common events/words to identify commonalities. A discussion of the methodology employed within this Stage is set out below, followed by an expansion on the methods employed.

The results of Stage 1 were used to create a *framework of controversies* which is presented at the end of this chapter[50]. This framework underpins the later tool design work undertaken in Chapter 6.

2.1 Methodology

As outlined in Chapter 1.1 above, the purpose of Stage 1 is the examination of a number of controversial STs; the aim being the identification and expansion of the social controversies which arise from the development and/or introduction of each ST. The most accessible sources of data on these controversial STs are in the form of documents, primarily consisting of media reports, journal articles, books, and reports by governments and NGOs. As such, to achieve the aim set here I have employed the methodology of a *multiple case study* based on documentary evidence. Given that I am examining multiple instances of a phenomenon (i.e., the social controversies associated with STs) at essentially the same point in time, this multiple case study approach can also been referred to as a *comparative analysis*, *comparative case study*, or *parallel case study* (Baxter 2010).

While there are multiple definitions for a case study, for the purposes of this project I employ the following; 'the intensive study of a single case or small number of cases where the purposes are to explain those cases and shed light on a larger class of cases' (incorporating Baxter (2010) and Gerring (2004 & 2007)).

[50] See Chapter 2.15.

Stake (2008) categorises case studies into three types; *intrinsic*, *instrumental*, and *multiple*. *Intrinsic* case studies focus solely on a single interesting case with no purpose of understanding some wider phenomenon; i.e., examining that individual case is the end justifying the means. *Instrumental* case studies describe where "*a particular case is examined mainly to provide insight into an issue…. The case is of secondary importance; it plays a supporting role, and it facilitates our understanding of something else*" (p.120). *Multiple* case studies describe an instrumental study which has been expanded to incorporate several cases "*chosen because it is believed that understanding them will lead to better understanding, and perhaps better theorizing, about a still larger collection of cases*" (p.123). I have chosen to employ *multiple* case studies as:

i. I am interested in a specific issue (that being the sources of social controversy within STs) rather than developing a deep understanding of a ST (thus ruling out intrinsic case studies). The narrowly-focussed examination of each ST serves to provide the necessary insights to understand the social controversies inherent to that ST.
ii. As my focus is on identifying whether the same controversies appear across STs, such that generic and universally applicable (as opposed to bespoke) design tools may become a possibility, it is necessary to incorporate several cases; hence the multiple case studies approach was decided upon.

The nature of the research undertaken within each individual case study is predominantly *cross-sectional* as opposed to *longitudinal*, in that it was undertaken at a single point in time[51]. There are two limited exceptions here. In relation to whole body scanners[52], regulatory decisions of 2013 to restrict the use of the backscatter variant within airports require a limited revisit, as do the conclusion of criminal court proceedings in 2013 over the development of the fraudulent ADE-651 Handheld Explosives Detector.

The case-studies display elements of both depth-orientated *idiographic* research and breadth-orientated *nomothetic* research (see Baxter 2010), without being committed

[51] The twelve incorporated case studies were undertaken throughout 2011.
[52] See Chapter 2.2.

to either method. They are idiographic to the extent that they seek to develop an in-detail understanding of that which provokes controversy in relation to the design and operation of ST; however they ignore all other aspects of those STs. Similarly, they are nomothetic in that they investigate the common nature of these elements as they appear across twelve different controversial STs; while at the same time ignoring STs which have not evoked social controversy.

Navigating this path between idiographic and nomothetic research is instrumental in my attempts to produce mid-range theories, as previously discussed in Chapter 1.7. Furthermore, by adopting this comparative approach I can counter a criticism sometimes levelled at case studies; that they "*focus too much on the particular and not enough on what is common across case studies*" (Baxter 2010, p.93).

As stated above, data sources for these controversial STs are *documentary*, in the form of media reports, journal articles, etc. Documentation is one of the most commonly used sources of evidence when conducting case studies. They benefit from being stable, unobtrusive, exact, and broad in scope; though it must be noted that they can also be difficult to retrieve, subject to selection bias, and may themselves be biased (Yin 2009).

For the purposes of Stage 1, multiple case studies were chosen over alternative qualitative data collection methods such as surveys, questionnaires, and interviews for three reasons. Firstly the multiple case study approach provides an established methodology for examining a particular phenomenon across a relatively small sample of subjects while facilitating the development of theories for application to the wider population. Secondly they can be undertaken using existing evidence (i.e., documents).

The third reason is related to the fact that individuals are simultaneously members of many different sub-groups within a society; one of these being those who consider a specific ST to be unacceptable. There is not an easily identifiable or accessible single sub-group of society who are against *each specific* ST, let alone *all* STs, making their inclusion as respondents for possible interviews or questionnaires unfeasible. Also the purpose here is not to identity common characteristics across those who oppose all (or the majority of) STs, rather to identify possible common sources of controversy within the STs themselves. Furthermore, those who oppose one of the STs incorporated as a

case study may indeed support all of the others STs included. Thus the problem of identifying suitable respondents will need to be addressed at the start of each separate case study undertaken (i.e., twelve times for twelve case studies). This point turns the conducting of interviews, questionnaires or surveys into an untenable option.

2.1.1 Multiple case study method

Before commonalities of controversy can be identified across different STs, it is necessary to individually examine each selected technology. This is achieved by undertaking individual document-centric case studies on the following twelve STs:

- Whole body scanners
- Profiling technologies
- Data matching
- Hand-held explosive detectors
- The UK's national identity scheme
- National identity cards
- Mosquitos
- Data mining
- Closed circuit television
- Less-lethal weapons
- The UK's national identity register
- Mass biometric systems

The purposes of each review are to identity and expand upon social controversies which arise from the development and/or introduction of that individual ST. This is achieved by narrowing the focus of the literature searches to isolate:

a. *Controversies arising from how the technology is designed*: By 'how the technology is designed' I am referring to; the physical construction of the technology, its innate characteristics, the scientific principles employed by and/or incorporated within the technology, and its functionalities and capabilities. To provide examples; for the design of a CCTV system this includes the resolution of the images captured, the ability to pan and zoom, the recording functions, whether it operates as a platform for other technologies such as face-matching, automated threat detection (ATD), and/or automated number plate recognition (APNR), etc. For an airport whole body scanner this includes the images produced, the incorporation of any privacy enhancing technologies (PETs), whether or not

ionising radiation is employed, what the dose rate is and the health effects of this, the time to complete one scan, what materials can be detected, etc[53].

b. *Controversies arising from how the technology is operated*: Aside from the design of a technology, a ST may evoke controversy by its method of operation. The controversy here may be *systematic* (i.e., if it refers to the normal usage of the technology - such as the injuries caused by kinetic less-lethal weapons[54], and/or usage as a result of an official policy – such as the refusal to all passengers to opt for alternative secondary screening procedures rather than undergo an airport body scan[55]) or *isolated* (i.e., in the event of specific instances of misuse/abuse of a ST by an operator – such as an individual CCTV operator using a street-based camera to spy through the windows of a private dwelling[56]).

c. *Negative public reactions to the concept of a particular ST*: This category acts as a back-stop to avoid excluding those social controversies not included within the design and operation of a particular ST. Such instances are likely to occur when a particular ST is being proposed by a government but has not yet been developed and/or implemented. An example here being national identity cards within a society traditionally opposed to such devices[57].

It must be stressed that social controversies readily cross the artificial boundaries created by these three categories; for example the image created by a backscatter body scanner[58] evokes elements of controversy falling within all three of the above categories. The purpose at this stage is to ensure a comprehensive scope of social controversies is incorporated into the individual case study literary reviews; not to seek to categorise these controversies into discrete boxes.

Non-social controversies are excluded from these case studies; by which I refer to controversies that do not manifest within the public domain and hence possess the capacity to evoke a social impact. For example this includes technical disagreements between engineers/scientists over a technology (or aspect thereof) within

[53] See Chapter 2.2.3
[54] See Chapter 2.13.3
[55] See Chapter 2.2.3
[56] See Chapter 2.10.3
[57] See Chapter 2.5.3
[58] See Figure 2.1

predominantly closed-source peer-reviewed journals; especially where such discussions are not reported and repeated within the wider public domain, and/or where social aspects of the technology are not discussed.

The sources included within these case studies are academic papers, newspaper articles, published reports (produced by governments, intergovernmental agencies, NGOs, and advocacy groups), legal cases, books, and online publications. Publications referring to controversies are predominately focussed on a single ST. To date there does not exist a corpus of work on the controversies created by STs *where the focus is on comparing and contrasting those arising across disparate technologies*, hence there are no seminal texts with a cross-technology focus. There are however respected writers who comment on different STs (such as Bruce Schneier), and those whose work has application across different STs (such as the PET work of Ann Cavoukian). These are included where appropriate.

The individual case studies undertaken do not represent an exhaustive historical account of all STs that have evoked controversy. Those included are limited to the twelve listed above. Their selection for inclusion is the result of a number of factors:

1. They represent a contemporary snapshot of the types of STs being introduced and/or expanded upon in the 10 years directly preceding the commencement of this research project[59]. Many have been linked to the *fight against terrorism* and have been suggested as appropriate security responses to the methods of terrorist attacks inclusive of, and subsequent to, the Al-Qaeda airline suicide attacks of September 11th 2001 within the US (hereafter referred to as '9/11').
2. These technologies are at the epicentre of controversies within different western societies.
3. The types of controversies they are linked to resonates at a societal level, with different societal sub-sets adversely affected by the design and/or application of these technologies seeking to express their opposition/concerns. It should be noted here there is no requirement that such protesting groups represent a majority of a society for a technology to be included.

[59] As a general guide I am focussing on the period from 2000 onwards. Many of the STs included within this paper were developed and implemented before this period, but where possible and appropriate I focus on their designs and methods of implementation from 2000 onwards.

4. All of the technologies listed above are the subject of negative press. As such the controversies are at least afforded the opportunity to be brought into the consciousness of the wider public.

The final selection of twelve STs is assisted by keyword searches[60] of individual UK newspaper databases. On the basis of these results this process of keyword searching is then expanded and repeated to include UCL library databases and Google Scholar for each identified ST to identify case study materials. This secondary process resulted in the identification of approximately 180 documents.

Ultimately the time constraints of this research project preclude a historically exhaustive review of all security technologies that have elicited social opposition. Regardless, given the broad scope of those STs included herein, this artificial truncation is not considered inappropriate or detrimental.

2.1.2 Coding method

The coding of the completed individual ST case studies entailed a three-step process:

Step 1: Starting with the first ST (whole body scanners; Chapter 2.2), and focussing on the identified associated problems with this technology (see Chapter 2.2.2), each individual instance or aspect of specifiable controversy pertaining to that ST within its literature review is identified. These are individually separated, listed, and assigned an identifier code (e.g. WBS1, WBS2 … etc.) within the individual tables titled *Identified controversies arising from n* (e.g. see Chapter 2.2.4) [61]. This step is repeated for each of the remaining eleven identified technologies; the tabulated results of which are presented in Chapters 2.**3**.4, 2.**4**.4, 2.**5**.4 … 2.**13**.4.

Step 2: The individual instances/aspects of controversy identified for all twelve STs in Step 1 are further categorised within Step 2 by both their *origin* and *nature*; the results of which were presented within Appendix *E* (see also Chapter 2.14).

The *Origin of Controversy* refers to 'where the locus of the controversy appears to originate'; i.e. does it appear to stem from:

[60] Including; *security*, *controversy*, *terrorism*, *society*, and *technology*.
[61] Substitute *n* for the name for each ST as they appear in Chapter 2.2-2.13

a) *Design Features*: aspects of the physical construction of the ST itself.
b) *End Users*: actions by, or criticisms by, the end-users of the ST.
c) *Policy Decisions*: political and/or institutional decisions; further divided into:
    i. *rules governing use*: specified codes-of-practice, best-practice, rules-of-engagement, rules for use as designated through primary or secondary legislation, etc.
    ii. *wider policy decisions*: higher level political and/or institutional policy decisions, which often form the basis/impetus justifying the lower level rules governing use.

The *Nature of Controversy* refers to eight broad categories which encapsulate the broad nature of each of the coded instances/aspects of specifiable controversy identified within Step 1 above. These are; health, legality, the public, rights & liberties, cost, safety & security, functionality, and use & misuse[62].

More than one *origin* and *nature* of controversy can be identified respectively when coding each controversy, with no upper-limit set.

Step 3: Finally for each coded instance/aspect of controversy *keywords*, *key-phrases*, and *key-concepts* were identified. These represented a reductionist attempt at expressing each of the controversies through usually no more than one or two words. Again more than one key-word/phrase/concept can be identified for each coded controversy.

2.2 Whole Body Scanners

To date *whole body scanners* (WBSs) have predominantly been utilised at airport pre-boarding security checkpoints. They operate by producing an image of an individual passenger which highlights any undeclared items concealed within or under that passengers clothing. There are two types of WBSs operating within airports, each based on different technologies; *backscatter X-ray scanners* (BXSs) and *millimetre wave scanners* (MWSs).

[62] Each is discussed in more detail in Chapters 2.14 & 2.15 below, before their incorporation into Table 1.13

BXSs conduct a 7-8 second high speed scan of a person's body with a narrow beam of low intensity X-rays as they stand motionless, either facing a single scanner (thus requiring a second scan to obtain *front* and *back* images) or by standing between two scanners which allows the simultaneous production of both images. The radiation backscattered (reflected) near the surface of the skin is measured by detectors within the scanner and converted into an image of the individual. This is displayed on a remote viewing monitor, identifying both metallic and non-metallic objects concealed within and under the individual's clothing (DHS 2009; NCRP 2003; Klitou 2008; Sweet 2009).

MWSs create an image by rotating antennas around a person emitting millimetre wave radio frequency energy. While both metals and the human body are highly reflective of these waves, ceramics, plastics and other organic matter (including organic explosives) are less reflective (Elias 2010). By collecting energy reflected off the body, a three-dimensional image is created that highlights any items being carried (TSA no date).

2.2.1 Purported benefits of whole body scanners

The purported technological strengths of incorporating WBSs into the current security regime include the following:

- It adds an extra layer of defence and detection to existing security measures; representing what Schneier (2006) refers to as *defence in depth*.
- WBSs possess the ability to identify metallic *and* non-metallic objects, plastic and liquid explosives (European Commission 2010), in addition to drugs, money, flora and fauna; all of which are smuggled under the clothing of airline passengers.
- Undertaking a whole body scan is quicker than undertaking a thorough pat-down.
- WBSs can reduce the negative human factors[63] inherent to the process of screening for the hidden objects listed above (Klitou 2008).

[63] *Negative human factors* entail "*the demands a job places on the capabilities of, and the constraints it imposes on, the people doing it. For screeners, the human factors issues cited in past studies include the repetitive tasks screeners perform, the close and constant monitoring required to spot the rare*

Arguments supporting the introduction of WBSs (again predominantly arising from the aviation security sector) are as follows:

1. WBSs enforce the passenger's right to security/safety

It has been the UK government's position to date that:

> Ultimately the rights of individuals must be balanced against the need to protect passengers and others at risk from terrorist threats and accordingly the use of [WBSs] … is, we believe, proportionate in these circumstances" (DfT 2010b, para.42).

2. As attackers adapt their attacks so must defenders adapt their defences

The crime scientist Paul Ekblom posits that "*[c]rime prevention faces a perpetual struggle to keep up with changing opportunities for crime and adaptable offenders*", likening this challenge to evolutionary struggles and military arms races (1999, p.27). The development of WBSs to prevent suicide bombers from smuggling liquid and plastic explosives into aircraft cabins provides a case in point.

3. WBSs are less intrusive than pat-downs

In a study of the publics' attitudes towards WBSs it was found that the vast majority of passengers[64] find the process of undertaking a WBS both less intrusive than, and preferable to, the traditional pat-down with its intrinsic physical contact (Mitchener-Nissen et al 2010).

2.2.2 Associated problems with the design of whole body scanners

Reflecting the substantial level of controversy surrounding the introduction and on-going operation of WBSs, a diverse variety of arguments are put forward challenging this security technology. These are set out below.

---

appearances of dangerous objects, and the stress involved in dealing with the public, who may dislike being screened or demand faster action to avoid missing their flights" (Klitou 2008, p.318).

[64] A ratio of 11 to 1 in favour of WBSs over pat-downs.

1. Legality concerns:

The legality of this technology is an issue which to date has never been addressed by a UK court or the European Court of Human Rights (ECtHR)[65]. Neither is there bespoke UK legislation specifically governing the presence and operation of WBSs. The question of legality can therefore only be addressed by examining pre-existing legislation that might be applicable to this technology. Any assumptions or conclusions made by this process will necessarily be contestable.

Commonly cited pieces of legislation which may be engaged by WBSs include the following:

- *Protection of Children Act 1978* (PCA): in relation to the images of children produced by these scanners. Early BXS trials did not include under-18s out of concern over potential breaches of the PCA (Telegraph 2009), specifically s.1(1)(a) *Indecent photographs of children*. This position changed after the publishing of the DfT's Interim Code of Practice which requires children undergo a scan if selected (DfT, 2010b); a position reinforced in the subsequent consultation paper which states "we will be requiring all children who are selected to be screened using the scanners" (DfT 2010c: p.9).
- *Human Rights Act 1998* (HRA): in relation to the potential contravention of the privacy rights under A.8(1) HRA. Based on previous ECtHR judgments the Court has adopted a wide view of what constitutes a person's private life thereby involving A.8(1)[66]; a concept which includes a person's right to their image[67]. Hence the images produced by WBSs are considered to fall within the existing legal ambit of a person's private life under A.8(1) HRA (Mountfield and Gearty 2010). The true question becomes whether the interference of this right by airport operators is justified under A.8(2) by possessing a legitimate aim, is proportional, and is in accordance with the law?
- *Equality legislation*: Finally there is the issue of compliance with the general equality provision which infuses UK law as made explicit in a number of statutes. It

[65] Within the US, the Federal Court of Appeals ruled in 2011 that WBSs did not constitute an unreasonable search (Kravets 2011).
[66] In *S and Marper v United Kingdom, nos.30562/04 & 30566/04 ECHR 2008*.
[67] As per *Sciacca v Italy, no.50774/99 ECHR 2005.*

has been argued that the Secretary of State for Transport failed to exercise his duty to ensure equality (as per the Race Relations Act 1976, Sex Discrimination Act 1975 and Disability Discrimination Act 1995) either when considering the introduction of WBSs into UK airports or when compiling the Code of Practice for their use (Mountfield and Gearty 2010).

2. Disability discrimination concerns:

In a submitted response to the body scanner consultation paper, the Disabled Persons Transport Advisory Committee (DPTAC) raise specific concerns regarding what they see as the *one-size-fits-all* approach adopted by the Department for Transport (DfT) whereby "*if a passenger is selected for security scanning, they will not be offered an alternative method of screening*" (DfT, 2010c para.21). In response DPTAC submit that "*there are certain disabilities with associated conditions and/or equipment where a physical search is preferable and where a security scanner would be either inadvisable, inappropriate and in some cases impossible*" (DPTAC 2010, para.4).

Because of the strict requirement of offering no alternative screening method, the DfT feels compelled to explicitly state that "*[p]assengers must not be selected on a basis that may constitute discrimination*" (2010c para.21). However in Manchester Airport, body scanners are operated in conjunction with traditional metal detectors, such that if the passenger triggers the metal detector they are required to submit to a body scan. As DPTAC points out, disabled passengers are more likely to require mobility equipment, assistance devices and medical devices which will trigger metal detectors thereby requiring the passenger be body-scanned. As such these passengers may experience indirect selection discrimination via the operational processes adopted by airport operators, despite the DfTs best intentions.

3. Equality and Human Rights Commission Concerns:

The Equality and Human Rights Commission (EHRC), in response to the DfT WBS Code of Practice consultation, outlines four concerns regarding the operation of WBSs (EHRC 2010).

a. WBSs infringe Article 8 of the European Convention of Human Rights.

b. The failure to publish the selection criteria employed for deciding who is scanned and the absence of independent monitors to police this operation opens the operation of WBSs to charges of discrimination and arbitrary usage.

c. The proposed future Code of Practice lacks sufficient details and consistency and as such may result in discrimination and arbitrary usage of WBSs.

d. The lack of evidence on either the effectiveness of WBSs or the positive impacts of their usage is a source of real concern.

It is the EHRCs view that "*there is a serious risk that implementation of body scanners will occur in a way that will discriminate directly or indirectly on the grounds of race or sex, in particular, and that their use will have an adverse effect on community relations*" (EHRC 2010, para.5).

4. Operational concerns relating to WBSs:

Since WBSs have been introduced into airports a number of concerns have arisen in response to the way these machines are operated *in the real world* by the airport administrators. Those that I have identified are discussed briefly below:

a. *Questions regarding the safety of undertaking WBSs*

Much of the debate surround the use of WBSs has focussed on the potential negative health implications of undertaking BXS scans. Contrary to the name 'backscatter scanner' these involve subjecting the passenger to a dose of ionising radiation as some X-ray photons penetrate the subject's body (Callerame 2010). Ionising radiation can result in the formation of mutated cancerous cells (Shapiro 2002), though the risk of developing a fatal cancer from one scan is approximately 1 in 1,000,000,000 (Mitchener-Nissen 2010; DfT 2010a) given the low dose rates involved.

According to the International Commission on Radiological Protection (ICRP), any measure which exposes individuals to ionising radiation must adhere to the principles of radiation safety (Ball and Moore 1997), including the *Principle of Justification*; whereby any exposure to radiation should do more good than harm (ICRP 2007).

There are two opposing positions as to whether BXSs meet this requirement; both arrived at by comparing different criteria. Klitou (2008) takes the view that the

justification requirement is not met for there is no positive net benefit as the intrusive nature of the images produced is disproportionate given the threat to security posed by liquid/plastic explosives, ceramic knives, etc., which he considers over-exaggerated. Strom (2005) however adopts the opposite position, claiming BXSs meet the justification requirement by comparing the increased security benefit they provide against the health risks they pose which he describes as so trivial as to be meaningless.

Despite repeated claims by various national governments' and the manufactures of BXSs that this particular technology is safe, in 2011 the European Commission adopted new guidelines for the use of body scanners within airports which effectively banned those scanners employing x-ray technologies (i.e., BXSs) (European Commission 2011). Alternative MWSs remained deployable.

b. *Controversies arising from the actions of the operators*

Various actions of scanner operators have resulted in negative media reporting of WBSs which may act to indirectly undermine public support for these scanners by reducing trust in the operators. These events include:

- A Heathrow Airport screener receiving a police caution for making lewd comments about the breasts of a female colleague after she passed through a scanner (Mirror, 2010).
- A screener in the United States attacked a colleague over taunts at work resulting from a WBS image taken of the screener during a training exercise which appeared to show he had a small penis (AoL News 2010; SMH 2010).
- An incident whereby 35,000 images were retained from a MWS in a Florida court building and subsequently posted on the internet (EESC 2011).

c. *Refusal to allow alternative screening methods*

One of the recurring criticisms of the operation of WBSs was the refusal within Europe to allow passengers the option of refusing a body scan, undertaking instead a physical pat-down. This was in contrast to the situation in the United States whereby passengers *must* be allowed to refuse a scan in favour of a pat-down because of their constitutional rights. Ultimately this social criticism within Europe culminated in the

European Commission adopting a new legal framework for security scanners in line with the US model whereby passengers have the right to opt out of a scan in favour of some alternative screening method (European Commission 2011).

d. *Children undertaking WBSs*

The controversies over children undertaking WBSs centre on the health risks of scanning children and of the legal status of the images produced[68]. The health concerns arose despite DfT claims that scans pose no additional threats to children (or a foetus) beyond that faced by adults (DfT 2010a).

e. *Use of profiling to select passengers for screening*.

Profiling is addressed in depth as a separate *controversial crime technology* in Chapter 2.7. For the purpose of the operation of WBSs in airports the main concern is that profiling will equate to discrimination under a different guise. This concern is reflected in the DfT Interim Code of Practice which states "*[p]assengers must not be selected on the basis of personal characteristics (i.e. on a basis that may constitute discrimination such as gender, age, race or ethnic origin*" (DfT 2010b, p.5). However the human rights group *Liberty* have criticised the Interim Code for failing to disclose what selection criteria are used to selecting passengers for screening; information withheld on national security grounds (Liberty 2010).

f. *The inability of WBSs to detect all concealed items*

When detecting non-metallic items, WBSs are not perfect. In reference to the failed airline attack by Umar Farouk Abdulmutallab, former Home Secretary Alan Johnson confirmed "*there would have been a 50 to 60 per cent chance of [the 3 ounces of PETN hidden in Abdulmutallab's underwear] being detected [by the WBSs]*" (Hansard 2010, col.34).

[68] The issue of the legality of the images produced by BXSs was addressed in detail in the Legality section above.

5. The intrusive nature of the images produced – specifically those by BXSs

Perhaps the single issue with the greatest propensity for provoking controversy is that of the graphic nature of the images produced by BXSs (see Figure 2.1). Like any image, whether or not an individual finds it to be pornographic, or a disproportionate breach of privacy, is a subjective judgement for that individual. Critics of these images are strong in their condemnation, equating them to *digital strip searches*. As Klitou illustrates "*the use of a backscatter body scanner, without the employment of a privacy algorithm, is comparable to conducting a strip search, and thus is considerably more intrusive than an appropriately conducted pat-down*" (2008, p.317). After years of complaints over the intrusive nature of these images and the subsequent failure by Rapiscan to modify their algorithms so as to create less revealing generic images, in 2013 the U.S. Transportation Security Administration cancelled their contract with Rapiscan. All 174 BXSs were removed from all US airports by mid-May 2013, replaced with less intrusive MWSs (Bloomberg 2013; New York Times 2013; Guardian 2013; Daily Mail 2013a).

Original image removed for copyright reasons from this electronic version.

Image can be viewed online at:

http://www.dailymail.co.uk/news/article-1240193/Body-scanner-wouldnt-foiled-syringe-bomber-says-MP-worked-new-machines.html

Figure 2.1 Rapiscan Secure 1000 scanner image

2.2.3 Reactions and responses to whole body scanner problems

The efforts to address WBS concerns have focussed on *operational fixes*, *technological fixes*, and *restrictive legislation*. The first seeks to implement the best procedures for how the scanning process is undertaken within an airport. The second focuses on developing the technology such that the scanners themselves are less controversial. While the third has resulted in BXSs being forcibly removed from airports in countries throughout the world.

1. Operational fixes

The primary operational fix has been the introduction of the Interim Code of Practice governing how airport operators are to operate their scanners. This code contains numerous procedures and requirements designed to minimise the propensity for controversy. A selection of these includes:

- Only security vetted and trained airport staff will be able to view the images produced by the WBSs.
- Security staff viewing images will be physically separated from the person being scanned.
- Security staff with the passenger will not be able to see the scan.
- All images are deleted immediately after analysis.
- The machines have no capacity to save, print or store the images viewed by the screener (DfT 2010b).

2. Technological fixes

Two avenues of technological fixes are currently being developed; (1) *privacy enhancing technologies* (PETs) whereby the images created by scanners are modified such that privacy is ensured without affecting the security offered by the scanner, and (2) *automated threat detection* (ATD) systems which remove the need for humans to view images by automating the process of detecting threats from the images created.

PETs and ATD are sometimes treated as *silver bullets* by academics, officials and governing bodies for *both* protecting privacy *and* ensuring security when utilising WBSs. Law professor Jeffery Rosen fêtes the substitution of computer generated images for those of the scanned individual as "*guarantee[ing] just as much security while also protecting privacy*" (2007, p.292).

Ann Cavoukian, the Information and Privacy Commissioner of Ontario, writes extensively on her concept of *privacy by design*: the "*approach of embedding privacy into the design specifications of various technologies*" (no date: p.1). Applied to WBSs, Cavoukian concludes "*[WBS] technologies that incorporate strong privacy filters – de-identifying raw images for backroom screeners, and using generic body images (or*

rendering body images to mere outlines) for frontline screeners, can deliver privacy-protective security" (2009a, p.6).

Additionally, a European Commission's communication to the European Parliament specifically referred to PETs and ATD as possible solutions for data protection, human dignity and fundamental rights concerns, achieved by:

- the computerised modification of images to protect privacy and prevent identification,
- the prevention of image storage, printing and transfer, and
- the eventual phasing out of human image interpreters as automated threat recognition technologies improve (European Commission 2010).

However these technological fixes do not always meet with immediate success. Within the United States, the Transportation Security Administration (TSA 2010) considered the current generation of ATD as insufficient to meet its unpublished detection standards. However, by 2013 this position had obviously changed as it had become a requirement that all airport body-scanners implement ATD systems so as to protect privacy.

3. Restrictive legislation

As discussed throughout Chapter 2.2.2 above, BXSs have been removed from all US airports on the basis of privacy concerns, as the manufacturer Rapiscan had been unable to integrate privacy enhancing ATD technologies into their products. To the same effect, but for different justifications, BXSs have also been removed from all EU airports on the basis that the radiation produced jeopardised citizens' health and safety.

2.2.4 Identified controversies arising from whole body scanners

Table 2.1: Identified controversies arising from whole body scanners

Code	Technology = **Whole Body Scanners**
WBS1	Potential breach of s.1 Protection of Children Act 1978; specifically the prohibition on creating indecent pseudo-photographs of children[69].
WBS2	Potential breach of Art.8(1) Human Rights Act 1998 & European Convention on Human Rights; the right to respect for one's private and family life. The images created by WBSs engage this right, so the question becomes whether the interference complies with the qualifications under Art.8(2); potentially the *proportionality* requirement, but more likely the *accordance with the law* requirement.
WBS3	Potential non-compliance with the *rule of law*; specifically the lack of clarity and sufficient narrowness of scope within the Interim Code of Practice[70] governing WBS usage to prevent arbitrary decision making. Specific concerns on this point include:
WBS3a	• Lack of a statutory basis for WBSs in the UK
WBS3b	• Lack of justification for exceeding EU standards
WBS3c	• Basis for selecting passengers not outlined within the Interim Code
WBS3d	• Few details governing the use of WBSs found within the Interim Code
WBS3e	• No independent review process, independent monitoring mechanism, or independent complaints mechanism
WBS3f	• No statutory scheme to safeguard against arbitrary selection processes whose effect can equate to discrimination based on religious dress, nationality, nation of origin, destination
WBS4	Potential non-compliance with the Race Relations Act 1976 and Sex Discrimination Act 1975; specifically the equality requirements under both Acts by (a) disproportionately inhibiting free movement by particular groups, and (b) by the lack of clear safeguards to prevent discrimination. Potential areas of discrimination here include sex, religion, religious dress, nationality, nation of origin, and destination.
WBS5	Potential non-compliance with the Disability Discrimination Act 1995; under the Interim Code of Practice no passenger selected for WBS screening will

[69] It is questioned whether the defence under s.1B(1)(a) applies whereby making the pseudo-photograph is permitted providing it is necessary for the prevention, detection or investigation of crimes.

[70] UK Department for Transport (2010) *Interim Code of Practice for the Acceptable Use of Advanced Imaging Technology (Body Scanners) in an Aviation Security Environment*.

	be offered an alternative screening method. The Disabled Persons Transport Advisory Committee noted that for certain conditions and/or equipment a WBS would be unadvisable, inappropriate, or impossible. Also this may constitute indirect discrimination as disable passengers are more likely to require mobility devices which will activate the metal detector thus requiring they be scanned by the WBS.
WBS6	Safety concerns arising from the use of X-rays within Backscatter X-Ray Scanners and the associated radiation dose rates received by those scanned; under exposure requirements of the International Commission on Radiological Protection's *principle of justification* any exposure to radiation should do more good than harm. If the harm caused by the intrusive nature of the images or the radiation exposure is disproportionate to the security threat posed by non-metallic materials then this may not be met.
WBS6a	• Specific health concerns over exposure risks to children
WBS6b	• Specific health concerns over exposure risks to pregnant women and their foetus
WBS6c	• Failure to release official specifications and dosimetry data
WBS7	Incidents of misuse of WBSs by airport staff undermining public support and trust; lewd comments by screeners and one screener attacking another over penis-size slurs.
WBS8	Posting of scanner images on the internet despite constant official assurances these machines have no capacity to store images; WBS public support is partly based on the understanding that these machines cannot store images, however 35,000 images from a Florida court WBS were posted on the internet.
WBS9	UK continues to refuse to allow pat-downs as an alternative screening method unlike the US where this choice is protected by constitutional rights.
WBS10	Failure of governments and WBS manufacturers to release statistical data proving the security effectiveness of WBSs; as a result of governments and manufacturers failing to release data on the ability of WBS to detect non-metallic materials, or data comparing WBSs to pat-downs to prove the effectiveness of WBSs, the public cannot make informed decisions on whether to support the introduction of WBSs. Also the case is not made justifying giving up a measure of privacy for a suitable increase in security.
WBS11	Passenger profiling concerns; the Department for Transport claims passenger profiling is not used to select passengers for WBSs however they do admit that in practice passengers are selected in response to *evidence-based concerns about a passenger*, which is profiling by another name.
WBS11a	• The selection criteria used to select passengers publically disclosed.

WBS12	Concerns over the graphic nature of the images produced by WBSs; these have been considered a breach of privacy and compared to *conducting a digital strip search*. Critics consider them more intrusive than a pat-down.

2.3 The UK's National Identity Scheme

On March 30, 2006 the Identity Cards Act 2006 (ICA) entered the UK statute book despite its choppy progress through both Houses of Parliament in the face of concerted opposition from within all political parties (Morris 2007-2008), public activist groups[71], and independent academic researchers. Whitley and Hosein (2010) describe the ICA as comprising seven components:

1. *the **National Identity Register (NIR)***: a central population register of every UK citizen and resident aged from 16 years;
2. *National Identity Registration Number (NIRNo)*: everyone in the NIR would receive a unique number to identify them;
3. ***[mass] biometrics***: everyone in the NIR would submit to various biometric sampling and recording, including at least fingerprinting;
4. *the **ID Card***: a physical card containing information from the NIR;
5. *legal obligations*: legal requirements to produce your ID card;
6. *administrative convergence*: the NIR and NIRNo would be used by agencies and organisations as an administrative base;
7. *legal liabilities*: the ICA created a number of crimes and financial sanctions to enforce compliance.

Collectively these components form the UK's National Identity Scheme (NIS). Chapter 2.3 focuses on the overall NIS while the following three Chapters 2.4, 2.5, and 2.6 focuses on national identity registers, national identity cards, and mass biometric systems respectively.

[71] For example: NO2ID.

2.3.1 Justifying the national identity scheme

A number of benefits have been put forward by those advocating for a national identity scheme. These include:

- Helping to protect cardholders against identity theft and fraud;
- Providing a reliable way for checking the identity of people in positions of trust;
- Make travelling within Europe easier;
- Creating a secure way of applying for financial products/services as well as conducting financial transactions (including internet based ones);
- Providing a simple and secure way of proving your age;
- Confirming eligibility for public services and benefits as well as reducing the amount of fraud relating to these benefits/services;
- Assisting in the prevention of organised crime and terrorism;
- Reducing illegal immigration into the UK as well as helping combat illegal working;
- Making it quicker for police to identify suspects, those incapacitated and those they have arrested (Morris 2007-2008; Whitley and Hosein 2010).

Additionally there are the arguments that implementing a NIDS will make it harder for terrorists to obtain fraudulent IDs (Morris 2007-2008), and that as states must already hold information on its citizens so as to operate an effective government then a national identity scheme will simply allow such information to be more effectively compiled. This will help ensure wrongdoers such as illegal immigrants and those engaged in criminal acts will not fall through the cracks of bureaucracy or be anonymous when travelling (Roy 2005).

2.3.2 Associated problems with the design of the national identity scheme

Prior to the ICA receiving Royal Ascent, in 2005 the *Identity Project* group within the London School of Economics and Political Science (LSE) published a comprehensive report assessing the then Identity Card Bill. It was the view of the authors that "*the establishment of a secure national identity system has the potential to create*

significant, though limited, benefits for society" (LSE 2005, p.9). Hence the authors did not reject outright the value of a national identity system. However, upon examining what the Government envisioned the UK's system would entail they conclude "*the proposals currently being considered by Parliament are neither safe nor appropriate … the proposals are too complex, technically unsafe, overly prescriptive and lack a foundation of public trust and confidence* (author's emphasis)" (LSE 2005, p.9).

1. The lack of public trust

For a NIDS to be successful it needs to be trusted by the public. This trust must encompass all aspects of the scheme including the technologies involved, the intentions of the state which oversees the NIDS, as well as how the NIDS operates and interacts with individual citizens.

It is held that the Home Office failed to address this public trust issue which resulted in a steady decrease of public support for the ICA NIDS, a general view that the scheme did not respect individual privacy, and the belief that it was "*relying too heavily on centralized government management of data*" (Whitley and Hosein 2010, p.210). Exaggerated government claims as to the benefits of the UK NIDS, the use of *spin*, and some dubious accounting methods, all did nothing to help their cause when attempting to win over the general public.

Finally there were trust concerns over NIDS *function creep* with Gordon Brown reported as planning "*a massive expansion of the ID cards project that would widen surveillance of everyday life by allowing high-street businesses to share confidential information with police databases*" (Hinsliff 2006).

2. Privacy and other civil liberty concerns regarding the NIDS

One area of considerable concern is the infringements of civil liberties inherent in the ICA NIDS. Primarily centred on privacy this also incorporates questions of informational self-determination and the propensity for future abuse.

While there are other national identity schemes operating throughout the world[72] there are three design features which distinguishes the UK NIDS from all others:

[72] For example the US social security number system.

1. The accumulation of a lifetime *audit trail* under s.1(5) ICA of every occasion an individual's identity is verified and information from the database is disclosed.
2. The construction of a biometrics scheme for an entire population to be used for broad purposes.
3. The insistence of a single *gold standard* of identity in order to generate trust, with the effect of either replacing or reframing the UKs social and economic relationships. (LSE 2005)

Concerns over this scheme included the following specific issues:

- Such pervasive surveillance could have a chilling effect on the actions of citizens (Gutterman 1988).
- The UK NIDS enrolment process entailed mandatory *biographical* checks whereby masses of data on the individual would need to be collected prior to their enrolment interview so as to form a series of questions for establishing the individual's identity (Whitley and Hosein 2010). This raises many privacy concerns, not least being the inherently intrusive nature of such a search, as well as questions over what happens to this data once the interview is successfully completed.
- This scheme entailed a lack of *informational self-determination* which is less about anonymity as about the ability of an individual to maintain control over what information is known about them (Roy 2005). Under the ICA NIDS the individual would have no direct control over who could access their information or what information was stored on the NIR.
- The NIDS offers a ready-made police-state tool of control for a future less trustworthy government.
- Demanding to see an individual's ID card could form a convenient pretext for those in authority to question and harass individuals on the basis of their appearance or ethnicity thus exacerbating societal divisions (NO2ID no year).

3. Questions over whether the ICA NIDS would achieve its goals in a reliable manner

On achieving its goals, the scheme envisioned creating a single secure biometric ID card which would be used for an incredibly diverse range of activities; from accessing

medical treatment, receiving state benefits, and opening bank accounts through to accessing pan-European travel hubs. The problems identified with linking all these benefits under a single card is the unintended consequences and lack of flexibility which arise from only having a single card and making this card indispensable (Whitley and Hosein 2010).

There are also serious doubts over the ability of the ICA NIDS to produce appreciable national security benefits by identifying terrorists. Terrorists would circumvent the proposed ID card requirements by (amongst other things) using tourist or student visas (as nineteen of the 9/11 attackers did), acquiring false identities outside the UK, spoofing identities and/or failing to carry a card which does not have to be produced on demand (Roy 2005; Whitley and Hosein 2010).

Additionally the reliability of the UK NIDS would depend on the quality of the initial biographical check conducted pre the initial enrolment interview. This quality is doubtful given the time, cost and capacity constraints in collecting reliable data on 50 million people. Furthermore, over time the number of errors within the data will steadily increase as a proportion of the total, as a result of errors in the entries; the classic *garbage-in garbage-out* problem which affects all data-sets (Roy 2005).

4. Public opinion ultimately turned against the NIDS

Despite reportedly high initial public support for the UK NIDS in 2002/3 (Home Office 2003) once information about the scheme's details started filtering through to the public, according to Morris (2007-2008) this support began to dwindle away as people realised what the scheme involved and how much it would cost. By July 2006 an ICM poll placed support for a UK ID card at 47% in favour to 51% against (ICM 2006). YouGov Daily Telegraph polls placed support for the NIR database at 22% happy - 78% unhappy, with only 11% of respondents trusting the government to keep the collected data confidential.

Two specific factors of the NIDS which undermined public support were the *compulsion aspects* (e.g. mandatory attendance for interviews and biometric sampling backed up by possible fines) and the *centralised database* as opposed to a federated scheme (Joinson et al. 2006).

5. The scheme was too complex to be delivered

The UK has a poor track history when developing previous large-scale information and communication technology projects (Beynen-Davies 2006); a challenge exacerbated within the UK's NIDS with its diverse range of objectives. Neither the *lifetime audit trail*, population-wide biometric system, nor the formation of a single ID aimed at transforming the nation's economic and social relationships, are present in any of the world's other NIDSs. It has been noted that such a format goes against the best-practice experience gathered from other countries experiences (LSE 2005).

On the operational side, it has been claimed the government was inaccurate in a number of their predictions which would have had fundamental repercussions. This included an overestimation of the usable life of stored biometrics, and an underestimation of the required high standard (and associated costs) of the biometric equipment required to both capture and compare biometric data over an entire population (LSE 2005).

6. The NIDS was itself a security risk

Those countries that have created a single general reference source for their citizens[73] suffer much higher instances of identity theft than the UK, as such nominally secure and trusted identification is much more useful to fraudsters (NO2ID no year). Additionally single national identity databases are valuable and vulnerable targets for attackers (whether they be they organised criminal gangs or foreign powers) and open to abuse by domestic forces. They have repeatedly been used to facilitate atrocities including genocide and ethnic cleansing; examples being the Nazi national ID document J-stamp; the racial information on ID cards under South Africa's apartheid system; and the ethnic classification information on the Rwandan ID cards which facilitated Hutu militia identify and kill Tutsis at roadblocks (Roy 2005).

7. The burdens and risks are placed on the citizen with the with benefits for the government

The argument was made that the scheme provided greater benefits for the UK government (e.g. assisting police, saving money through public service administration,

[73] Such as the USA and Australia.

etc.) than it did for citizens (Whitley and Hosein 2010), while at the same time the risks posed by small errors within the scheme's administration fell on the citizens. NO2ID (no date) argued that by making large swathes of ordinary life dependent upon the reliability of a complex administration system, small errors within this information could have potentially catastrophic repercussions for an individual by denying them their rights to public services or their property with no immediate solution or redress.

8. The NIDS was too expensive

It was asserted that the cost of the NIDS would far exceed the Government estimates (approximately £5.84 billion over ten years), ending up somewhere between £10.6 billion and £19.27 billion over ten years. Furthermore, as biometric technologies are changing and improving all the time, unless the initial creation of the NIDS IT infrastructure was designed to be compatible with these changes, it will be increasingly difficult (and expensive) to update the future system infrastructure (Whitley and Hosein 2010).

2.3.3 Reactions and responses to the national identity scheme problems

Following the change in government after the 2010 general election, the incoming Conservative and Liberal Democrat coalition made good on their respective election manifestos to repeal the ICA. To do so the coalition passed the Identity Documents Act 2010 that repealed the ICA on 21 January 2011, signalling the end of the NIS in its current form. Under s.2 Identity Document Act 2010 (see Appendix *C*) all ID cards were invalidated on this day. The physical manifestation of the NIR was destroyed on 10 February 2011 with the mechanical shredding of 500 hard drives. Conservative MP Damian Green, who participated in this physical act, was quoted as saying "*[w]hat we are destroying today is the last elements of the national identity register, which was always the most objectionable part of the [ICA] scheme*" (Mathieson 2011).

Despite both (a) the much-vaunted security applications of the NIR and ID card system, and (b) the money spent developing this technology from a policy initiative to a realised entity, that this wholly premeditated act of destruction was met with no public outcry is itself a telling fact.

2.3.4 Identified controversies arising from the national identity scheme

Table 2.2: Identified controversies arising from national identity schemes

Code	Technology = **National Identity Scheme**
NIS1	Exaggerated claims over outcomes, spin, and dubious accounting; all undermined public trust and support in the scheme when they came to light, and force the government to make embarrassing admission.
NIS2	Concerns over function creep; plans to expand the NIDS project such that high-street businesses would share confidential information with police databases and surveillance would be widened to include the actions of everyday life, such as using a cash card or an iris scan machine to enter a building.
NIS2a	• Doubts over government's ability to resist pressures to use collected data for purposes beyond that for which it was collected given the cost of the scheme.
NIS3	Privacy concerns over the NIDS;
NIS3a	• Questions about *informational self-determination* given personal data was collected mandatorily and used for purposes outside of the individual's control.
NIS3b	• Information on an individual within the NIDS is collected, retrieved, and processed without the individual's consent or even their knowledge; hence the individual effectively has no control over the information about themselves.
NIS3c	• Concerns over the potential for future abuse.
NIS4	The lifetime audit trail; would record every instance an individual's identify is verified and every time information on an individual held on the database is disclosed.
NIS4a	• This data would form a record of a person's life covering such mundane activities as purchasing goods and withdrawing money and would be retained even after the death of the individual.
NIS4b	• The presence of this audit trail would have a chilling effect on the actions of individuals.
NIS4c	• The systematic observation of individuals by the government negates an individual's efforts to maintain private information about where they go, what they do, and who they associate with.
NIS5	The *biographical check* component; required mass data be collected on every

NIS5a *NIS5b*	individual prior their enrolment interview. • Raised privacy questions as to the intrusive nature of searches required to collect this data. • Questions over what happens to the data afterwards, how is it stored, who has access to it, what protections are in place, who is it shared with, how are the interviewers vetted, etc.
NIS6	The NIDS offers a ready-made tool for controlling citizens within a police state under a future less trustworthy government.
NIS7	The Home Secretary had the power to designate new classifications under which to register individuals which would permit abuse and discrimination.
NIS8 *NIS8a*	Not certain NIDS would enhance security; terrorists are trained to *blend-in* thus avoiding actions which would make them noticeable under NIDS. • Terrorists/criminals could circumvent NIDS requirements by using tourist visas, student visas, acquiring false identities outside the UK, spoofing identities, failing to carry a ID card, etc.
NIS9 *NIS9a*	Garbage-in-garbage-out; the time, cost and capacity restraints of collecting information on 50 million people for the initial checks raises doubts as to the quality of the system it will produce. • Errors within the system will steadily increase as the data entries increase, making analysis of this data increasingly prone to error.
NIS10	Overinflated public support through question framing and sampling at a time of crisis.
NIS11 *NIS11a* *NIS11b* *NIS11c* *NIS11d* *NIS11e*	People withdrew support from the scheme in response to; • Public awareness as to what the scheme entailed. • The compulsory aspects of the scheme (i.e. mandatory interviews, biometric sampling, financial sanctions, etc.). • A centralised database and lack of public trust in governments to secure information on databases following a number of embarrassing failures. • The cost. • The drop in perceived threat levels.
NIS12 *NIS12a*	NIDS is a security risk; placing such valuable data on a single database creates a honey-pot for attackers. • Countries with single universal reference sources for citizens suffer higher levels of identity theft as the nominally secure and trusted ID source is given greater weight by other citizens/companies thus is

	more useful to fraudsters.
NIS13	Single NID databases have been used repeatedly to facilitate atrocities; such as the Nazis, SA apartheid, Rwandan ethnic cleansing, etc.
NIS14 *NIS14a*	Scheme benefitted the government more than it benefitted citizens. • The risks posed by small errors in data having potentially catastrophic effects fell on the citizens, potentially denying them the right to public services and their property with no immediate solution or form of redress.
NIS15	The scheme was too expensive.

2.4 The National Identity Register (NIR)

The NIR was a component of the Identity Cards Act 2006 (ICA) and comprised a central population register of every UK citizen and resident aged from 16 years. The purposes of the NIR were set out in s.1(3) ICA[74]; namely the creation of a method for individuals to prove their identity and to collect together facts on individuals which could be ascertained and verified when necessary and in the public interest to do so. Section 1(4) defined the scope of what was 'necessary in the public interest', while s.1(5)(i) created the 'audit trail' requirement which would record every occasion an individual's NIR information is provided to another. Every individual on the register would also be assigned a unique number known as their National Identify Registration Number (NIRNo).

2.4.1 Purported benefits of the national identity register

The purported benefits of the NIR seem only to be limited by the imagination of those developing policies and/or technologies which require the presence of (or would be enhanced by) a system for the collation of identifying information on every citizen within the UK. A very small snapshot of these benefits incudes:

- the provision of a system for proving people's identity;

[74] See Appendix *B*.

- the creation of an audit trail detailing whenever an individual's identity is verified against the register;
- when coupled with the NINRo it facilitates the creation of a system whereby information could be shared between both different government departments/agencies and with private organisations;
- enhancing national security; and
- reducing identity fraud.

2.4.2 Associated problems with the design of the national identity register

Cost of the NIR

Introducing a new uniform numbering system for the entire UK population will require all government agencies and departments modify their individual systems to operate on this identifier. As Whitely and Hosein (2010) point out:

> [a]t present over 80 departments and agencies have their own unique identifiers for each record because that numbering system is appropriate for their systems, processes, and policies. Introducing a new uniform numbering system will not only be costly but also burdensome and most likely unnecessary. (p.111)

Privacy concerns

There are serious privacy concerns over creating a single all-encompassing government database allowing the surveillance of citizens that would also be accessible to private companies and individuals, such as insurers, landlords and employers (Morris 2007-2008). Because of the scope of information collected on the NIR, the Information Commissioner Richard Thomas publically attacked the scheme questioning both this scope and the wide range of bodies to which this information would be made available. He was not alone in these concerns with the House of Lords/House of Commons Joint Committee on Human Rights expressing their own deep concerns over the individual privacy violations represented by the NIR as a component of the ICA.

Disproportionate collection of information under the audit requirement

The systematic collection of information of individuals under s.1 ICA (especially the audit trail requirement under s.1(5)(i)) could amount to the entire life history of an individual which would be retained even after their death (Morris 2007-2008; NO2ID no year). Not only would citizens not be able to opt out of this recording process they would be legally required to notify the register of any changes in their registerable facts regardless of their personal beliefs as to the legitimacy of the NIR.

Security concerns

Of all the concerns regarding the NIR it is probably the security issues such a database would create which attracted the most critical attention. Various security critiques of the NIR have identified the following issues:

- Because of the scale and complexity of the NIR it is infeasible to build a computer system that can provide the necessary level of security assurance required to protect the safety and privacy of records contained therein (LSE 2005).
- There is no obligation on the Secretary of State to protect the NIR data or to authenticate an individual's consent permitting their records be accessed (LSE 2005).
- Police and others will have almost unconstrained access to data held in the NIR, opening UK citizens to serious new risks and creating significant new opportunities for criminals (LSE 2005).
- A centralised singular ID database creates an inherent and serious security risk, acting as a *honey-pot* for hackers, those engaged in identity theft, attacks by foreign governments, and insider attacks. This threat is exacerbated by the lack of an explicit obligation within the ICA to ensure data held in the NIR is secured with appropriate access controls (Joinson et al 2006; LSE 2005). Furthermore the NIR database would need to be on a network if it is going to be usable by border checkpoints, police, and others. This will enhance its vulnerability to hackers (Roy 2005).

Abuse of data held in the NIR by third parties and insiders was a recurring complaint by critics. On the issue of third party abuse, NO2ID (no year, p.6) noted the following problems created by the ICA:

> The requirement that all those registered notify all changes in details risks creating the means of tracking and persecution through improper use of the database. A variety of persons have good reason to conceal their identity and whereabouts; for example: those fleeing domestic abuse; victims of "honour" crimes; witnesses in criminal cases; those at risk of kidnapping; undercover investigators; refugees from oppressive regimes overseas; those pursued by the press; those who may be terrorist targets. The seizure of ID cards (like benefit-books and passports now) will become a means for extortion by gangsters.

On the specific issue of insider attacks, given the volume of data collected and processed daily by the NIR (estimated at 1 gigabyte of data per day), analysing the data to detect suspicious behaviour by users/operators becomes a massive problem. This would not be adequately tackled through automated/semi-automated audit analysis programs, rather would also require manual audits which are costly and subject to budget cuts (LSE 2005).

Legal concerns

There are three areas of specific legal concern regarding the formation and operation of the NIR:

1. Article 8 HRA – Right to respect for private and family life

Under Art.8 the ICA obviously constitutes a breach of privacy rights. However this right is not absolute and can legitimately be breached by the UK government providing (amongst other things) it is *necessary in a democratic society*. For any infringement to by the state to be lawful, courts have required such actions be both *proportionate* and *necessary*. On the proportionality issue Morris (2007-2008) argues the range of information collected by the ICA is so broad that it will exceed its stated goals. On the *necessity* requirement Khan (2006, pp.142-3) questions whether the ICA is indeed necessary to meet legitimate aims for a number of reasons, stating:

> It seems that the Government has failed to prove the case for such a scheme for several reasons. First, terrorist atrocities do not always stem as a result of false identity. Recent calamitous attacks in London have shown that suicide bombers do not always worry about the concealment of identity. Secondly, the use of biometric devices could have significant health implications in that the safety of

such devices has not been fully tested. Thirdly, the immigration status of an employee can be amply determined by the employer through a perusal of the employee's passport or National Insurance number. Fourthly the project is expensive and difficult to monitor…. Finally, the scheme could be controversial in circumstances where personal details on the central database can be accessed by public sector organisations without the individual's consent.

2. Data protection principles

As the NIR would hold an increasingly large amount of data on all those registered, and this data is matched to each individual (creating the *digital footprint* of that person) data protection concerns are engaged, specifically in relation to the follow principles:

- *Purpose* (personal data shall be obtained only for one or more specified and lawful purposes, and shall not be further processed in any manner incompatible with that purpose or those purposes) - Both the broad nature of the uses outlined in s.1 ICA and the intention to link many different public service providers to the scheme may undermine this requirement.
- *Proportionality* (personal data shall be adequate, relevant and not excessive in relation to the purpose or purposes for which they are processed) - The NID infrastructure raises concerns over data creep should it become an all-purpose system for monitoring and controlling the UK population. Also the proposed NIR system theoretically allows for behavioural profiles to be created for every member of the UK population.
- Security – The creation of a central registry containing personal identification data for the UK population raises serious security concerns given the value of this information to criminals. As such "*[a]ppropriate security measures, technical and organisational, should be taken by data controllers to protect personal data from unintended or unauthorised disclosure, destruction or modification*" (Beynon-Davies 2006, p.15). The technical challenges of achieving this security would also be considerable.

3. The rule of law

Finally for an Act to be considered in accordance with the *rule of law* the requirements therein must be sufficiently clear and foreseeable so as to place individuals on nation of its application to them, and thus avoid constituting arbitrary interference with the

citizen. Morris (2007-2008) questions whether the ICA meets this requirement on the basis that as information can be entered onto the NIR without the consent or knowledge of the individual then they may be unaware their privacy right is being breached.

2.4.3 Reactions and responses to the national identity register problems

Reactions and responses to NIR problems are the same as those already covered in relation to the NIS. See Chapter 2.3.3 above for this discussion.

2.4.4 Identified controversies arising from the national identity register

Table 2.3: Identified controversies arising from national identity registers

Code	Technology = **National Identity Register**
NIR1	Creating a single numbering system used by over 80 government departments/agencies for all UK citizens;
NIR1a	• Would be very costly.
NIR1b	• Would force these departments/agencies to modify their own bespoke systems designed specifically to suit the needs of these bodies.
NIR1c	• Would be a burdensome process, and most likely unnecessary.
NIR2	Privacy concerns over creating a single NIR; would be used for citizen surveillance and accessible by private companies/individuals such as insurers, landlords, and employers.
NIR3	The systematic collection of information under the audit requirement would amount to a person's life history, retained after their death. It could not be opted out of, and citizens would be legally required to register and information of any changes to registrable facts.
NIR3a	• Many people have legitimate reasons for concealing their identity and whereabouts which would be undermined by the requirement they must keep the ID database updated.
NIR3b	• This includes those fleeing domestic violence, victims of honour crimes, witnesses in criminal cases, those at risk of kidnapping, stalker victims, refugees from oppressive overseas regimes, terrorist targets, etc.

NIR4	A computer system cannot provide the necessary level of security assurance to protect the privacy of the information given the scale and complexity of the NIR.
NIR5	No obligation on the Secretary of State to protect NIR data or to confirm consent by an individual when their records are accessed.
NIR6	Police and others will have almost unconstrained access to data held on the NIR, creating serious new risks for citizens and opportunities for criminals.
NIR7	Centralised singular ID database is the perfect target for hackers, inside attackers, or foreign governments.
NIR7a	• No explicit obligation to ensure appropriate access controls to the NIR.
NIR8	NIR will need to be on a network, enhancing its vulnerability to hackers.
NIR9	Vulnerable to insider attacks; the volume of data/daily processes means automated/semi-automated audit analysis programs will not adequately detect insider attackers.
NIR10	• Manual checks are expensive and will be prone to budget cuts, hence many insider attacks will probably go unchecked.
NIR11	NIR potentially breached Art.8(1) HRA;
NIR11a	• The information collected may not meet the *proportionality* requirement of Art.8(2) as the information collected is so broad that it will exceed its stated goals.
NIR11b	• The information collected may not meet the *necessity* requirement of Art.8(2) as;
NIR11b1	○ terrorist atrocities don't always result from false identities,
NIR11b2	○ suicide bombers don't always worry about concealing their identities,
NIR11b3	○ the immigration status of an employee can be adequately ascertained by passports and national insurance numbers without resorting to a NIR.
NIR11c	• Public sector organisations could assess an individual's data without their consent.
NIR12	Concerns regarding European data protection principles, specifically;
NIR12a	• *Purpose*; personal data under the NIR would be collected and processed for purposes other than those for which it is collected.
NIR12b	• *Proportionality*; collected data should be adequate, relevant and not excessive in relation to the purposes for collection.

NIR12c *NIR12d* *NIR12e*	• *Accuracy*; given the size of the NIR and the constant addition to data to it, inaccuracies will increase and affect the individual concerned. • *Anonymity*; personal identifiable data should be held only as long as required for the purpose for which it was collected. The NIR system allows behavioural profiles to be created for every UK citizen. • *Security*; appropriate security measures should be taken to protect data to prevent unintended or unauthorised disclosure. Given the value of the NIR to attackers and the range of people who could access it, maintaining security would be extremely difficult.
NIR13	NIR may breach the *rule of law*; it would not be sufficiently clear and foreseeable to citizens when their rights have been breached as people/organisations could unlawfully access their data without their knowledge.

2.5 National Identity Cards

This was not the first time NIRs or ID cards have been used in the UK. During both World War I & II national registers were created, with the WWII register combined with an ID card linked to wartime rationing. While the WWI register did not survive into peacetime, the WWII system lasted for a period after the war (Agar 2005; Whitley and Hosein 2010). While accepted by the public upon its introduction, discontentment levels rose after WWII ended, especially in response to the increasing practice of police demanding to see ID papers. As a direct result the National Registration Act 1939[75]. was allowed to lapse in 1952 (Roy 2005).

It is arguable that the UK Government ignored the lessons from its previous national identity schemes. While there seems to be support for national ID cards (or at least the concept of one) within the UK during times of crisis, such support appears to quickly wane once the threat justifying their introduction has passed or if unpalatable details of the scheme become widely known.

[75] The National Registration Act 1939 introduced at the beginning of World War II.

2.5.1 Purported benefits of national identity cards

National identity cards have been successfully introduced into other European countries and accepted by their citizens (Beynon-Davies 2006). Out of a total of twenty-two such schemes twelve are compulsory and ten are voluntary (Whitley and Hosein 2010). Indeed following 9/11 there have been public opinion polls showing strong support for national identity card schemes with both the UK and the US (Home Office 2003; Roy 2005); two countries which have traditionally opposed such schemes.

The UK Government's website for identity cards cited the following benefits as arising from such a scheme:

- ID cards will help in the protection of people from identity fraud and identity theft;
- They will strengthen security and improve public confidence;
- They will help tackle immigration abuse and illegal working;
- They will disrupt the use of false and/or multiple identities by those involved in terrorist activities and organised criminals;
    - They will ensure that only those who are entitled to do so will be able to use free public services (Whitley and Hosein 2010).

Indeed the former Home Secretary David Blunkett stated in 2004 that ID cards would help the fight against organised crime and terrorism, and stop people from using multiple identities (BBC News 2004). However it should be noted that by 2009 Blunkett had changed his position on ID cards and a compulsory national ID database holding it would be sufficient to introduce mandatory biometric passports instead (Guardian 2009).

Other supporters of national identity schemes have utilised utilitarian arguments holding that such schemes constitute a worthwhile trade-off in that we sacrifice a small reduction in anonymity for a significant gain in security. Though as Roy points out such positions have been based on the assumption that ID cards will be *fool-proof* and that and problems arising from the excessive collection and storage, or misuse of information held by a national identity scheme are separate issues in the earlier utilitarian quantification (Roy 2005).

2.5.2 Associated problems with the design of national identity cards

It is the view of Whitley and Hosein (2010) that while the national ID cards proposed under the ICA were viewed by the UK government as a simple solution to complex problems such as identity theft, fraud and terrorism, the government failed to realise that this security technology is itself a complex entity which creates its own complex problems. These controversies were such that the national ID card, along with the rest of the ICA, did not survive the change of government in 2010.

Firstly was the risk of increased crime that accompanies a national ID card to the extent that they actually make people less secure with one than without one. Being extremely valuable documents, there is greater incentive to forge national ID cards, as well as the fact that "*[n]o matter how unforgeable we make it, it will be forged*" (Schneier 2008, p.98). Indeed there has not been a card created that cannot be forged regardless of the security measures incorporated. An example being the French national ID card which was promoted as un-forgeable when introduced in the 1990s but forgeries of which now enjoy brisk trade. Even adding biometrics to the ID card does not guarantee security as iris scans, fingerprints and facial recognition systems can all be fooled (Roy 2005), and should the IC cards incorporate radio frequency identification (RFID) technology then there is the growing security risk that they will be read from ever increasing distances by attackers (Hunter 2005). Even if an attacker lacks the technical skills to forge the card itself, they can simply forge the documents required to obtain a legitimate card , or even more simply they could bribe one of the officials issuing the cards (Schneier 2008).

Creating hard to forge singular ID cards increases the risk of identity theft as well as the successful fraudulent use of the cards. If they are promoted as the gold standard in identification (as the UK ID card was) then it will be trusted for more and more applications under a linked, centralised system, thus allowing a fraudster to commit more offences with a single forged card than would be possible if many organisation issued their own bespoke cards (Schneier 2008). It will also require a centralised database which, as identified above, will be vulnerable to attacks by hackers and

insiders, data compatibility problems, erroneous and unreliable data as well as mistakes during data entry (Schneier 2008).

Moving beyond the security risks created by national ID cards there was the controversial question of whether they would actually be able to make us safer from terrorists, and there is strong evidence that this was not the case and that the UK government had come to realise this fact. To highlight, in 2004 then Home Secretary David Blunkett claimed UK ID cards would boost the fight against terrorism, though by 2009 he had reversed his position on this in the face of the mounting evidence from recent attacks in London and Madrid.

Firstly an ID card, while it may tell us something about the carriers identity, it tells us nothing about their criminal intentions (NO2ID no year; Schneier 2008) achieving little to nothing by way of prevention (per Lord Carlile in Morris 2007-2008) and so are effectively useless post an event. This is crucially important when the attackers a country is facing have no previous recorded links to terrorism or are legitimate citizens and therefore entitled to a national ID card; such as Timothy McVeigh, the DC snipers, many of the 9/11 bombers, the 7/7 London bombers and the Madrid bombers (Schneier 2008; Whitley and Hosein 2010).

Secondly research suggests no link exists between the prevalence of terrorism within a country and the presence of a national ID card, hence they do not significantly deter terrorists (NO2ID no year; Privacy International 2004) "nor are they recognised by analysts as a meaningful or significant component in anti-terrorism strategies" (Privacy International 2004, p1). Specifically it has been noted that:

Of the 25 countries … adversely affected by terrorism since 1986, eighty per cent have national identity cards, one third of which incorporate biometrics. This research was unable to uncover any instance where the presence of an identity card system in those countries was seen as a significant deterrent to terrorist activity. … [Indeed almost] two thirds of known terrorists operate under their true identity (Privacy International 2004, p.2).

Focussing on the scheme anticipated under the ICA, it is claimed that the UK ID card would only assist in anti-terrorism efforts if terrorists were (a) willing to register for

one, (b) used their true identity when doing so, and (c) there exists intelligence data that can be connected to that identity (Privacy International 2004).

Looking beyond terrorism concerns, traditionally there are other risks to citizens from ID cards, especially for ethnic minorities. Police harassment of minorities have been noted in countries with mandatory national ID cards, and such systems offer ready-made tools of oppression for future less-trustworthy governments (NO2ID no year).

Other purported benefits of a national ID card are also brought into question. Regarding illegal immigration and working, it is difficult to see how ID cards will prove any more an effective deterrent then passports or visas given that these requirements are already in place and yet ignored (NO2ID no year). Nor, it is claimed, would the ICA ID card model be effective in reducing identity fraud (LSE 2005).

Finally there is the issue of the lack of public support for ID cards, which was especially problematic given the recognition that both public trust and business buy-in were essential for the success of the ICA scheme (Whitley and Hosein 2010). This trust was undermined on a number of fronts including, (1) the excessive cost of the scheme, (2) the fact that the NID would fail to meet its (over)stated objectives, and (3) concerns over privacy rights and the excessive collection of data involved (Morris 2007-2008).

2.5.3 Reactions and responses to the national identity card problems

Reactions and responses to NID problems are the same as those already covered in relation to the NIS. See Chapter 2.3.3 above for this discussion.

2.5.4 Identified controversies arising from national identity cards

Table 2.4: Identified controversies arising from national identity cards

Code	Technology = **National Identity Card**
NIC1	NIDCs used for many official purposes are extremely valuable to attackers/criminals;
NIC1a	• Using an ID card for multiple purposes and promoting it as secure and un-forgeable makes it extremely valuable especially to attackers. The associated increased risk of crime from these attacks will actually

	make people less secure.
NIC1b	• Hard to forge singular ID cards increase the risk of identity theft and the successful fraudulent use of the card.
NIC1c	• Unjustifiably promoting a card as the *gold standard* in identity security, and *un-forgeable*, increases public/business trust in the cards such that they will automatically be considered genuine making the task of the fraudster even easier.
NIC1d	• Linking these cards to a centralised system rather than each department/business having their own bespoke card and system allows an attacker to make many more attacks with a single card.
NIC2	No ID card in history has ever been un-forgeable.
NIC2a	• The addition of biometrics will not make an ID card un-forgeable.
NIC3	Incorporate RFID technology adds the risk ID cards can be *read* from ever increasing distances by attackers.
NIC4	Security measures are easily circumvented; if the attackers lack the skill to forge the NIDC, they can forge the documents required to obtain a card, or simply bribe one of the officials issuing the card.
NIC5	A centralised database for NIDCs creates problems;
NIC5a	• It is a honey-pot for hackers and insiders, presenting a single target on which to focus their efforts.
NIC5b	• It compounds problems of data compatibility, erroneous and unreliable data, and mistakes during data entry.
NIC6	No evidence NIDC will make citizens safer from terrorists:
NIC6a	• NIDCs tell us something about the carriers *identity* but nothing about the criminal *intentions*
NIC6b	• Doesn't assist in crime prevention, and are effectively useless post an event.
NIC6c	• No causal evidence of a link between *prevalence of terrorism within a country* and *a NIDC within that country*, hence they do not significantly deter terrorists.
NIC7	Traditionally NIDCs pose a risk to ethnic minorities through police harassment; police have used supposed random checks of NIDCs as an excuse to detain, delay, and harass minorities.
NIC8	NIDCs provide a ready-made tool for oppression by future less trustworthy governments.
NIC9	Lack of public trust in the NIDC scheme due to excessive costs, privacy

	concerns, and the excessive collection of personal data, undermined the probability of the schemes success.

2.6 Mass Biometric Systems

By conflating natural physiography and bio-dynamics, a biometric can be defined as "a measurable physiological and/or behavioural trait that can be captured and subsequently compared with another instance at the time of verification" (Beynon-Davies 2006, p.7). This definition includes, amongst other things, fingerprints, ear-prints, blood-vessel mapping, iris scans, hand and knee geometry, face and voice recognition, handwriting and signature matching, keystroke patterns, and DNA matching.

The ICA envisioned recording various biometrics, such as fingerprints, iris scans and facial scans, from all those enrolled in the NIR, in effect creating a national biometric database. This data could also be stored on the national ID card and would be used for authentication and identification purposes. According to Beynon-Davies (2006) such biometric systems operate over three stages; (1) the sampling stage where a biometric sample is collected from the subject; (2) the storage stage where this sample is transformed into a digital template (which is typically encrypted) for storage on a database and potentially a local token (like an ID card); and (3) the recognition stage whereby a biometric reader measures a subject's biometric and compares this against the stored template.

2.6.1 Purported benefits of, and justifications for, mass biometric systems

Proposed benefits of the large-scale use of biometrics include the following:

1. To enforce border control; both by incorporating biometrics into passports so as to tighten access controls at a country's border, and within the issuing of visas so as to reduce instances of over-staying (Goldstein et al 2008; Grijpink 2006).

2. To better regulate asylum applications and avoid instances of *asylum shopping*[76] (Goldstein et al 2008).
3. The automation of access controls and the management of social processes; both within private companies and for government services. According to Grijpink (2008 p.317) "*biometrics is the only way of physically determining to whom a document, object or piece of data relates*".
4. For law enforcement purposes. Particular benefits cited here include; tackling identity fraud, counteracting illegal immigration, enhancing crime-scene investigations, and preventing crime by identifying suspects in advance of criminal acts (Goldstein et al 2008; Grijpink 2006).

There is also the goal of tackling terrorism. According to Privacy International (2004) UK Ministers have suggested four ways national biometric ID systems may be used to combat terrorism:

1. *A central database of biometric identifiers will detect whether a person is using multiple identities.*
2. *A process of comprehensive "biographical footprint checking" will help determine whether a person is using a false identity.*
3. *A comprehensive vetting of card applicants might detect those people who have a background that is indicative of a terrorist profile.*
4. *The existence of a compulsory identify card will expose those terrorists in the UK who have not registered.*

2.6.2 Associated problems with the design of mass biometric systems

The effectiveness of biometrics for these purposes depends on the type of biometric chosen, the quality of the sample taken (which itself is affected by a multitude of factors including the skill of the person collecting the sample, the individual from whom the sample is taken, the quality of the machine used to capture the sample, and the environmental conditions where the sample is collected), and the methods used to

[76] The process whereby an asylum-seeker attempting to obtain asylum within the EU lodges an application with one Member State having already been rejected by another.

compare the individual's biometric with the earlier sample held on file (Whitley and Hosein 2010). Each of these factors can undermine the effectiveness of a biometric system, increasing the rates of false positives (false matches) and false negatives (false non-matches) during the recognition stage.

On the issue of the reliability of the different biometrics, each one has its own bespoke problems affecting a minority of the population, thus the total number of those affected will increase as the population of the system increases. For example fingerprints are affected by: people with missing digits, the elderly and those with dry skin, and those whose hands are calloused or suffer regular damage from their work (such as labourers, chefs, farmers, builders, etc.). Iris scans are affected by cataracts and those who suffer conditions such that they are unable to hold their head steady or fix their gaze on a single spot. Facial recognition systems are affected by beards, makeup and weight changes (Hunter 2005; LSE 2005; Whitley and Hosein 2010). Each of these groups is a minority but they may well find themselves disproportionately affected and disadvantaged by the fact they cannot properly enrol in the scheme nor be able to use the various scanners during the recognition stage if they are successfully enrolled onto the national database. They may suffer constant delay and embarrassment as they will always be singled out for secondary screening. Furthermore, the types of biometrics suggested for use under the ICA can be forged. Iris scans can be fooled using patterned contact lenses, fingerprints can be copied using 'gummy fingers', and US researchers have identified dozens of ways to fool facial recognition systems (Roy 2005; Whitley and Hosein 2010).

However the real problems with mass biometric systems are the rates of false positives and false negatives, especially when the systems attempt to move beyond one-to-one-matching *authentication* and try to deliver one-to-many-matching *identification*. To expand, all biometric systems suffer the problems of false positives and false negatives. One can reduce the propensity threshold for *either* of these two problems occurring during the recognition stage (for example by adjusting the tolerated rates of false-positives or false-negatives) by adjusting the level of match/miss-match at which an alarm is sounded. However this will have the concomitant effect of increasing the probability that the *other* problem (the rate of

false-negatives or false-positives) will occur. Should the rates of either of these problems occurring be too high then the system will soon lose the trust of those administering it as well as those subjected to it. These false matches may be manageable if we are operating a system of authentication whereby we are performing a one-to-one comparison[77] with an accuracy rate of 99.99%[78] whereby an individual will on average only trigger a false alarm once every 10,000 times their biometric ID is authenticated. However while biometrics is good at authentication it is bad at identification (which requires one-to-many checking) especially as the size of the database to be checked increases. For example under the proposed ICA scheme with eventually 50,000,000 people registered a face-recognition, iris scan or fingerprint system, with 99.99% accuracy would result in every person who is checked by a biometric scanner being falsely identified as a terrorist 5,000 times; and that is every single person every single time they are scanned. Such a system is useless, fails to provide any measurable security benefits, and would quickly be ignored (Roy 2005; Schneier 2006).

Finally there are two further problems with biometric systems. Firstly, all a biometric ID does is match a person to their card. It does nothing to ensure that the original information and documents used to obtain this card were genuine or legitimately obtained (Roy 2005). This goes back to the problem identified earlier that promoting biometric ID cards as the *gold standard* of identification measures will undermine security and increase fraud as they will be unquestioningly accepted as legitimate. Secondly there is the problems that once a biometric stored on a database has become compromised through identity theft it is very difficult to restore the victim's identity through the production of a new digital identity as they cannot spontaneously generate new biometrics for themselves (Roy 2005).

Ultimately the mass biometric system envisioned under the ICA required UK citizens put their trust in unproven technology with questionable robustness which would result in errors resulting in inconvenience (or worse) for these same citizens. The cost

[77] Comparing just the sample taken during the recognition stage with that taken during the initial sampling stage.

[78] 99.99% is the probability the system will correctly identify a match (a positive positive) or not-match (a positive negative) without producing either a false positive or false negative. This level of accuracy is well beyond the capabilities of any current commercial biometric system.

of creating this system (including the multitude of readers which would be required throughout the UK if the system is to be usable once created, and the training of people to operate these readers) would also be enormous. It is highly questionable whether the security benefits flowing from such a system would be worth this outlay.

2.6.3 Reactions and responses to mass biometric system problems

While the UK NIR and UK NID Cards containing biometric information have been scrapped for the time being[79], one form of biometric has been included on UK passports. Also known as ePassports, these machine-readable passports have been issued since 2006 and contain a digitised image of the holders' face allowing for facial recognition. While the previous Labour Government intended to introduce *second generation* ePassports in 2012 incorporating fingerprint data, this measure was rejected by the new Coalition Government as confirmed within their Coalition Agreement of May 2010. As such it has not been acted upon (Gower 2012).

2.6.4 Identified controversies arising from mass biometric systems

Table 2.5: Identified controversies arising from national/mass biometric systems

Code	Technology = **National/Mass Biometric System**
MB1 *MB1a* *MB1b* *MB1c* *MB1d* *MB1e*	The effectiveness of biometric systems is not a given, but depends on many variables including; • The type of biometric sampled. • The skill of the person taking the sample. • The quality of the machines used to capture the sample. • Environmental conditions. • The methods used to compare an individual's biometric with the earlier collected sample.
MB2	Fingerprints are affected by; people missing digits, the elderly, those with dry skin or other skin conditions, workers whose hands are regularly damaged or calloused.

[79] See Chapter 2.3.3 above for more detailed information.

MB3	Iris scans are affected by; cataracts, the blind, those with condition such that they cannot hold their head steady or fix their gaze on a single spot.
MB4	Facial recognition systems are affected by; beards, make-up, and weight changes.
MB5	Minority groups are adversely affected; such as certain age groups, certain occupations, and sufferers of certain medical conditions. They may all find themselves disproportionately affected by a scheme they either cannot enrol in, or where they constantly suffer delay and embarrassment as they are constantly singled out for secondary screening.
MB6	Biometrics can be forged; patterned contact lenses for iris scans, gummy fingers for fingerprints, and there are dozens of ways to fool facial recognition systems.
MB7	All biometric systems suffer from false positives and false negatives. Reducing the threshold for one with increase the likelihood of the other.
MB8	A one-to-many biometric system within 50,000,000 people (like the UK proposed one) with a 99.99% accuracy would still result in a person being falsely identified as 5,000 other people every single time they are scanned. Such systems are operationally useless and would quickly be ignored by security staff.
MB9	Biometric systems just link a person to a card. They cannot check that the original documents/information used to obtain this card were genuine, nor can they tell us anything about the intentions of the person even if this card was obtained legitimately.
MB10	Once a biometric stored on a database is compromised it is very difficult to restore a person's identity through producing a new digital identity as they cannot spontaneously create new biometrics for themselves.

2.7 Profiling Technologies

Profiling refers to both:

a) The process of building profiles; i.e. discovering correlations between data in databases that can be used to create a representation of an individual or group (referred to as a *profile* which is essentially sets of correlated data).

b) The process of trying to match a subject to a profile; i.e. applying profiles of individuals or groups to a subject to determine the likelihood this subject matches the previously created profile (Schermer 2011).

While the term *profiling* has other common uses in related fields; most notably *DNA profiling* and *criminal profiling* which is defined as "*the process of using available information about a crime and crime scene to compose a psychological portrait of the unknown perpetrator of the crime*" (Muller 2000, p.235). When I refer to profiling I will be restricting the concept to definitions 'a' and 'b' listed above.

2.7.1 Purported benefits of, and justifications for, profiling

The primary purported benefit of profiling is that by identifying those individuals who, as a result of matching your previously determined criminal profile, possess the highest probability of themselves being a criminal, you can better target your security resources by subjecting these individuals to additional security/screening measures. It is argued this will result in a net drop in successful attacks as well as manage costs by improving resource allocation.

Additionally it is argued that profiling acts as a deterrence, dissuading prospective criminals from certain activities and/or places out of fear of capture before they can achieve their goal.

2.7.2 Associated problems with the design of profiling

Determining the efficacy of profiling is incredibly difficult, especially if the algorithms which form profiles are kept secret under the auspices of intellectual property and/or national security. Without the publishing of automated profile hits, search results, arrest numbers and subsequent conviction statistics, along with the details of the profiles themselves, it remains impossible to form an accurate statistical picture of the effectiveness of automated profiling.

Indeed the feasibility of even creating a profile of certain criminals such as *the typical terrorist* is highly suspect. There are nearly as many variant personalities for terrorists

as there are variants of personality, and without an identified *terrorist personality* there is nothing to form the basis of a usable profiling algorithm that will identify terrorists while not *red flagging* most of the population (Hudson 1999). Terrorist organisations recruit people for operations who look *normal* and do not stand out. Because the physical and behavioural descriptions of terrorists could describe almost any normal young person, terrorist profiling based on personality, physical, or sociological traits would not appear to be particularly useful (Hudson 1999, p.63). This view is supported by MI5's behavioural science unit through a leaked document which recognised that terrorists in the UK do not match popular stereotypes (they are not sexually frustrated youths, religious zealots, loners, gullible, unintelligent or suffering mental illnesses) rather they "*are a diverse collection of individuals, fitting no single demographic profile, nor do they follow a typical pathway to violent extremism*" (Travis 2008). According to Borrion et al (2008) it is still virtually impossible to build a system that will detect unknown terrorists before an attack, and it not clear whether any future technology will succeed in this regard.

However, if we obtained statistical data that terrorists within a specific organisation predominantly shared certain racial, national, religious or cultural traits then a profiling system based on these could narrow the statistical probability that an individual attempting to board a plane is a terrorist (Hudson 1999)[80]. However by definition such a system of religious or racial/ethnic profiling would be inherently and unavoidably discriminatory in nature.

Focussing on racial profiling, Press (2010) powerfully sets out just how offensive such a measure is:

> Racial profiling, as commonly defined, is any actuarial method that conditions an individual's prior probability of criminal behaviour explicitly on his or her race, ethnicity, nationality or religion. Mature democratic societies recognise racial profiling as not merely another type of actuarial policing, but as something deeply corrosive of democratic values. Racial profiling violates the democratic covenant that individuals are to be judged by a universal rule of law, not by shifting standards that vary with their being assigned to stereotypical racial or other categories (p.165).

[80] However, the best that could be achieved here would be this narrowing of statistical probability rather than the ability to identify terrorists with any degree of regularity, as the statement 'all terrorists are Australians' does not equate to 'all Australians are terrorists'.

This challenges the position that profiling technologies are ethically neutral. Profiles, to be valid, must be based on data and where this data is itself infused with racial bias then any subsequently created profile (and its operation) will itself be discriminatory. Furthermore as noted by Schermer (2011), because profiles are created using only selected pieces of the puzzle that makes up an individual, those elements that are included in the profile will have their effects magnified. Thus racial bias (and discrimination) is exacerbated by these profiles. Given that Ministry of Justice figures show that black people were 10.7 times more likely, and Asian people 2.2 times more likely, to be stopped and searched than white people under discretionary police powers within the UK (Runnymede 2010) it is easy to see how crime figures can become biased against blacks and Asians, and if this becomes a variable within a profiling algorithm then the algorithm itself become subsequently tainted.

Further controversies arising from profiling include what Schermer (2011) refers to as the *de-individualisation* of individuals from within minority groups as the profiles judge them based on the fact that they share common characteristics and traits with the *group profile* while ignoring the characteristics and merits of the *individual* being compared. Group profiling can reinforce the stigmatisation and stereotyping of the group in question and damage social cohesion, especially when people start treating others based on their knowledge of the stereotypes and not on the individual that they see before them.

Taking de-individualisation further, another criticism of profiling is that an individual is probabilistically identified as a likely terrorist/criminal and subject to whatever security repercussions this may bring on the sole basis of their conformity with a pre-determined profile of the typical terrorist/criminal *a priori*. As such it is not based on any evidence of *actual* criminal conduct on the part of the individual (Press 2009).

Examining profiling from a purely utilitarian perspective it is claimed that profiling fails to provide a net benefit to society. This is especially the case when profiling is based on race where it is claimed that the moral toll this inflicts on a democratic society outweighs any security benefits which arise (Press 2010), or to put it more directly the costs far outstrip the benefits (Runnymede 2010).

It also deliberately imposes a greater burden on the liberties of a minority within society (as determined by the profile itself) so as to provide a security benefit to the majority. This burden is itself compounded by the fact that those chosen by a profile are not subjected to secondary screening as a false positive on just one occasion, rather they will be repeatedly subjected to this additional screening regardless of how many times they have proved their innocence in the past, as profiling operates via selection *with replacement* (Press 2010). Each time the same individual passes through a profiling security checkpoint such as an airport terminal they will continually be flagged as a potential threat if no mechanism exists for downgrading the threat quotient of that individual in response to their repeatedly displayed innocence, thus the burden imposed on these individuals becomes compounded.

2.8.3 Reactions and responses to profiling technology problems

Despite the concerns over profiling, this technology has found widespread and diverse use in the crime and security spheres. Responses restricting its usage have largely focussed on the factors used to form the profiles themselves. Given the enormous potential for profiling technologies to be misused as implements of discrimination, efforts have been made to prevent controversies from arising by controlling what variables can be used to form profiles.

In the Department for Transports' Interim Code of Practice for the Acceptable Use of Advanced Imaging Technology (Body Scanners) in an Aviation Security Environment, it is explicitly stated that the selection criteria for determining which passengers will undergo airport body-scans must not be based on their personal characteristics. Examples of such characteristics are listed as including; gender, age, race or ethnic origin (DfT 2010b). This reflects the variety of UK anti-discrimination legislation now drawn together within the Equality Act 2010 which would apply to profiling but which was enacted without any specific security technology focus. This includes discrimination based on age, disability, gender reassignment, marriage and civil

partnership, pregnancy and maternity, race, religion or belief, sex, or sexual orientation[81].

2.8.4 Identified controversies arising from profiling technologies

Table 2.6: Identified controversies arising from profiling technologies

Code	Technology = **Profiling Technologies**
PT1	Impossible task; there is no profile for a *typical* terrorist as there is nearly as many variant personalities for terrorists as there are variants of personality.
PT2 *PT2a*	Terrorists possess *normal* personalities and physical characteristics; • They are recruited specifically on the basis that they are indistinguishable people who look normal and don't stand out.
PT3	Automated systems cannot detect unknown terrorists before an attack; there are no terrorist-specific features from which we can form a profile to check against.
PT4	Profiles can be discriminatory; when based on statistical data employing racial, national, religious, or cultural traits.
PT5 *PT5a* *PT5b* *PT5c* *PT5d* *PT5e*	Racial profiling which that calculates an individual's prior probability of criminal behaviour by their race, ethnicity, or religion should be avoided because; • Racial profiling is corrosive to democratic values. • Violates the democratic convention that individuals are to be judged by a universal rule of law, not shifting standards based on racial stereotypes or categories. • Democracies require in practice more than simple majority rule but also an acceptance of the rights of minorities. • It is fundamentally objectionable to disadvantage anybody because of their race. • It can create resentment by those profiled, and leave victims feeling paranoid, depressed, vulnerable, and targeted.
PT6	Profiling technologies are not ethically neutral; racial bias in the data will infuse the profile.
PT7	Profiles magnify racial bias and discrimination; as only selected elements of

[81] These represent *protected characteristics* under S.4 Equality Act 2010.

	an individual are included, thus their effects are magnified.
PT8 *PT8a* *PT8b* *PT8c*	Group profiling has negative effects; • It ignores the characteristics and merits of individuals within that group. • It reinforces stereotypes. • It damages social cohesion, especially when people start treating others in accordance with the stereotype and ignore the individual they see before them.
PT9	Profiling is not based on any actual evidence of criminal conduct by a person; when it identifies individuals as terrorists/criminals based on probabilistic conformity to a profile.
PT10	Profiling does more harm than good when based on race; any immediate security benefits are outweighed by the moral toll this inflicts and the increased risk of protest and violence by those who feel disenfranchised and discriminated against.
PT11	Profiling minorities may inadvertently increase overall crime rates; when secondary screening resources are focused on a narrow proportion of the population who match a profile, the remaining majority of the population may increase their criminal activities on the basis that they are less likely to be stopped.
PT12	Profiling places a greater burden on the liberties of a minority of the population for the security benefit of the majority.
PT13 *PT13a* *PT13b*	Individual false positives are repeatedly targeted; the burden of profiling will be repeatedly imposed on the same innocent individuals as selection for secondary screening operates via selection *with replacement*. • Being cleared after one infliction of secondary screening does not mean that individual's risk profile is amended. • Because finite secondary screening resources are repeatedly applied to the same individuals who match the profile but are innocent, profiling is no more effective a selection method than randomised selection

2.8 Data Mining

Data-mining is the application of database technology and techniques (such as modelling and statistical analysis) to data to identify valid, novel, implicit and

potentially useful information and patterns within that data (Schermer 2011; Steinbock 2005; Tien 2004).

It originated in the commercial sector where it is routinely used to target advertising, identify customer buying patterns, assess creditworthiness, etc. Within the fields of crime and security, data-mining is employed with aims which include detecting fraud and other criminal activities and patterns, detecting terrorists and analysing intelligence (Steinbock 2005).

Data-mining can be distinguished into *descriptive* and *predictive*. Descriptive data-mining algorithms seek to discover previously unknown correlations between different data objects and their attributes within a database. By doing so one hopes to gain insights into the populations from which the objects were drawn.

Predictive data-mining seeks to "*make a prediction about events based on patterns that were determined using known information*" (Schermer 2011, p.46). When applied to profiling this entails mining-data of an individual to determine the probability that they match a previously established profile; such as the probability an airline passenger matches the previously constructed profile of a terrorist.

2.8.1 Purported benefits of, and justifications for, data mining

Automated data-mining has made possible the analysis of huge amounts of data to identify patterns and relationships which previously would have, (a) been hidden by the vast scale of the data available, or (b) would have remained implicit within each separate data source but is made explicit by combining and mining a variety of sources (Kreimer 2004-2005).

The *subject* of the data-mining process is incredibly flexible. It is not limited to a person, but could also be an email address, a place, event, telephone number or purchase (Tien 2004). The list here could be endless.

The patterns identified by data-mining can be used to build or strengthen profiles, thus allowing more targeted searches, and hopefully reducing the number of false positives/negatives. Also by identifying previously hidden patterns and relationships,

data-mining can assist both policy makers and those implementing policies to improve the quality of their decision taking.

Finally it can be argued that computerised decision-making brought about by data-mining (and data-matching) is ethically and inherently neutral in that "*a computer algorithm will undoubtedly produce the same decision every time, with no bias or favouritism toward any party*" (Steinbock 2005, p.45).

2.8.2 Associated problems with the design of data-mining

The controversial aspects of data-mining falls into a number of categories.

1. What data-mining represents:

a) Steinbock (2005) notes that data-mining is associated with both a fear of totalitarian-style state observation, as well as the targeting of individuals by governments. One previous data-mining project 'MATRIX' was designed to identify potential terrorists from the general population based on their statistical *terrorism quotient* (i.e. propensity to commit terrorism) by data-mining an individual and comparing the results against a pre-constructed terrorist profile. This system produced a list of 120,000 people identified as being statistically likely to be terrorists which was passed on to US authorities (Tien 2004). It is obvious is that if there were 120,000 terrorists operating in the US then the country would be under constant attack, therefore MATRIX either produced too many false positives or didn't work at all.
b) There is also the perceived threat of *information asymmetry* whereby data-mining may alter the *status quo* existing between citizens and their governments by handing to the government masses of data on every citizen. As will be discussed below, knowing that every action will be recorded by a watching government may have a chilling effect on the willingness of individuals to engage in political activities deemed critical of the government and making them subject to future official decisions without knowing why (or even that this has occurred) (Schermer 2011).

2. Operational concerns with data-mining:

a) Data-mining is used to make *proactive* administrative decisions which can have concrete negative implications for an individual in the absence of any initiating action and are virtually impossible to challenge due to the lack of transparency surrounding the decision making process (Steinbock 2005). It is also questionable how society can also effectively govern such searches (Tien 2004).

b) The fact that the algorithms which underpin the decision-making process are not revealed (for intellectual property and/or national security reasons) makes it hard to mount an effective legal challenge. If you do not know what evidence was used to make the decision affecting you then you do not know what it is you need to challenge (Steinbock 2005).

c) When data-mining results in some form of secondary screening (for example a strip search at an airport), it may be impossible to determine whether the data-mining algorithm is good enough to produce sufficient *reasonable suspicion* to justify these secondary measures (Steinbock 2005).

d) The negative repercussions of being a false-positive under a data-mining/data-matching algorithm will not be *one-offs* rather will reoccur every time that individual enters a situation where they are assessed against that same algorithm. This will be compounded further if the algorithm is shared between government departments or within the private sector, increasing the situations where innocent individuals will be wrongly identified as being whatever their profile identifies them as (Steinbock 2005).

e) Computer-based reasoning (based on a digital decision-making algorithm) is harder to both evaluate and understand than human assessments; whereby the assessor can set out their thought processes justifying their decision. Because of the number of variables that can make up a data-mining algorithm they become little more than *black-boxes*, and computer neural networks can give no rationale for why they made the decision that they did (Steinbock 2005; Tien 2004).

f) Just because a data-mining programme produces a correlation between an individual and the elements of a profile, this only represents a correlation

between variables. It does not represent causation; i.e. it does not explain *why* that individual possesses those characteristics or has performed the activities that they have done (Schermer 2011).

g) From a legal perspective, administrative decisions with potentially adverse consequences are being carried out in the complete absence of any form of due process. As Steinbock (2005, p.45) notes, "*[d]ata matching and data mining give no process as the law understands that term. There is no evidence, no opportunity to be heard, no confrontation with adverse evidence, and no reason given – only a result. Under any theory of due process, decisions based solely and irrevocably on the results of data matching or data mining are deficient, at least when they affect substantial interests*".

h) There is little by way of public accountability or oversight of data-mining (Schermer 2011) especially when private companies are involved.

i) Administrative decisions arising solely from data-mining lack the benefit of *human judgement*, with a human's ability to weigh evidence and apply their knowledge of the real world, the environmental situation and current events (all of which will be outside the scope of a static algorithm) to inform their decisions. Nor can algorithms be persuaded to change their output in the face of reasoned arguments or additional information; they are a black box lacking the ability to independently interact (Steinbock 2005).

j) Interoperability of databases and computer systems is critical to the success of data-mining projects seeking to join-up the data held by different government agencies. However the fact that these systems were not designed with interoperability in mind (indeed in some cases they have been designed not to be accessible/usable by outside agencies) can limit the usability, accuracy and efficacy of data-mining systems (Seifert 2007).

3. The quality and nature of the data

a) A US Congress report noted that "[d]ata quality is a multifaceted issue that represents one of the biggest challenges for data mining. Data quality refers to the accuracy and completeness of the data. Data quality can also be affected by the

structure and consistency of the data being analysed. The presence of duplicate records, the lack of data standards, the timeliness of updates, and human error can significantly impact the effectiveness of the more complex data mining techniques, which are sensitive to subtle differences that may exist in the data" (Seifert 2007, p.21).

b) The results from data-mining processes which examine a person's behaviour with the goal of predicting terrorist intent have been described by Jeff Jonas[82] as being "*so far from reaching the level of accuracy that's necessary that I see them as nothing but civil liberty infringement engines*" (Gellman et al 2006). However we are increasingly using the results from data-mining to make important decisions about the criminality of individuals even despite these shortcomings (Steinbock 2005).

c) In response to the claim that algorithms are ethically neutral, if the data being mined is biased (or simply inaccurate) or was collected using biased means then the results of the data-mining will also be tainted. Furthermore these algorithms are created by people, and in construction if they have based any of their decisions regarding what variables to include/exclude on the basis of biased perceptions (with or without malice) then again the algorithm is a tainted entity.

3. Civil liberty based arguments:

a) Data-mining processes have a *chilling effect* on political expression. The monitoring of innocent people may deter them from engaging in legitimate protests out of fear of repercussions. Similar tactics were employed in the US by Nixon and McCarthy in response to Vietnam War protests and the perceived threat of communism; a fact acknowledged by the US Defence Department (Kreimer 2004-2005).

b) Privacy controls were weakened in the aftermath of 9/11 where under the auspice of fighting terrorism companies surrendered massive amounts of data they had collected on customers/employees, etc., to the government (Kreimer 2004-2005). This represents *mission-creep*, whereby data is used for purposes other than which it was originally collected. The problem here is that such data, having been

[82] Chief scientist at the IBM Entity Analytics group.

originally collected for other purposes and by questionable means, is often lacking in reliability which can result in increased false-positives (Seifert 2007).

c) Data-mining of massed data from disparate data sources makes explicit that which was implicit (Kreimer 2004-2005) thus affecting privacy and opening the individual to exploitation.

d) Privacy helps protect people from being abused by others, especially those in power or possessing nefarious intent. As Kreimer (2004-2005) points out, it is hard for a government to retaliate against a dissenter when they cannot track who that dissenter is. Also Schneier (2006) writes that systems such as the now defunct *Total Information Awareness* programme would have produced databases that would themselves have been prime targets for criminals (both outsiders and trusted insiders).

e) There is refusal by the owners/operators of the algorithms to disclose exactly what variables are included within them as such disclosure will negate the effectiveness of the algorithms by enabling circumvention (Steinbock 2005). However this lack of transparency forever leaves the data-mining algorithms open to charges of racial, religious, ethnic, and other forms of discrimination. It also deprives the potential for feedback loops with members of the public who can identify problems with, or improvements to, the algorithm which would improve it (Muller 2000).

f) The practical repercussions of being a data-mining *false positive* includes the stigmatisation and embarrassment of the innocent as they are, sometimes publicly, forced to undergo secondary screening to prove their innocence having been deprived of their presumption thereof by the data-mining/data-matching processes (Steinbock 2005).

g) Data-mining is being used for *fishing expeditions* whereby authorities search *before* they suspect. Democratic societies which value the presumption of innocence should suspect first *before* the search (Tien 2004). Otherwise this is merely the digital equivalent of police randomly carrying out house to house searches without any prior information or reasonable suspicion before doing so.

2.8.3 Reactions and responses to data mining problems

Responses by governments to data-mining have been mixed to date. At least two US data-mining projects that have been shut down in response to public outcries based on the fear of overreaching state observation; these being the Terrorist Information Awareness[83] (TIA) programme and the Multistate Anti-Terrorism Information Exchange (MATRIX) programme. Both of these were intended to mine immense quantities of data, potentially on an entire population. However, through revelations by whistle-blower Edward Snowden, we now know that another mass electronic surveillance data mining programme entitled PRISM has been secretly developed (Greenwald and MacAskill 2013) which potentially may be exceed the scope of either TIA or MATRIX[84].

In 2007 a US Bill requiring federal agencies to report to Congress on any activity to use or develop data mining to identify terrorist or criminal activity[85] and the impacts this would have on citizens' rights died in Committee. And yet outside of *security-centric* data-mining, in 2013 Massachusetts introduced a Bill banning the mining of student data for commercial purposes (Wired 2013).

[83] Originally designated the Total Information Awareness programme.

[84] The controversy over PRISM is ongoing, having divided politicians both within the United States and between the United States and other national governments. PRISM has not been included within my dissertation as; (a) information on its existence was only made public in June 2013 (see Greenwald and MacAskill 2013) and hence near the end of the write-up of my dissertation, and (b) as I stated within the methodology section of Chapter 2.1, the case-studies undertaken within my dissertation were cross-sectional in nature and not longitudinal. Hence new candidates were not included after these analyses were completed. While undoubtedly qualifying as a controversial ST, it remains to be seen what effects (if any) the currently expressions of public resistance will have on the design and operation of PRISM over the mid- to long-term.

[85] Entitled 'A bill to require reports to Congress on Federal agency use of data mining', this would have required annual reports be made available to the public covering, amongst other things; (1) a description of data-mining activities and goals, (2) an assessment of the efficacy of mining and the impact on individual privacy and civil liberties, (3) a discussion of the policies and procedures to protect individual privacy and due process and to guard against harmful consequences of potential inaccuracies. This information was taken from www.govtrack.us

2.8.4 Identified controversies arising from data mining

Table 2.7: Identified controversies arising from data mining

Code	Technology = **Data Mining**
DMI1	Data mining is associated with the fear of totalitarian-style state observation and the targeting of individuals by governments.
DMI1a	• These fears have resulted in two data mining projects being shut down in the US.
DMI2	Data mining which produces too many false positives will be ignored by security staff.
DMI3	Intrusive data mining systems represent a disproportional trade-off of individual privacy and liberty for minimal security benefits.
DMI4	Data mining resulting in proactive administrative decisions without any initiating action by the individual can have negative implications for that person.
DMI5	Impossible for individuals to effectively challenge the legality of decision making based on data mining;
DMI5a	• Lack of transparency surrounding the decision making process.
DMI5b	• When the algorithms used are not revealed as you do not know what evidence was used to make the original decision so you cannot produce counter-evidence to refute the decision makers evidence
DMI6	Society's governance of data mining is questionable;
DMI6a	• Complex mining algorithms and opaque decisions make society's governance questionable.
DMI6b	• Little public accountability or oversight of data mining, especially when conducted by private companies.
DMI6c	• Without disclosing the algorithm used, data mining is open to charges of racial, religious, ethnic, and other forms of discrimination. This approach is not accepted in other aspects of the criminal law (i.e., forensic evidence, financial evidence, etc.).
DMI7	It is nearly impossible to challenge the reasoning of the decision maker when it is a computerised decision making algorithm;
DMI7a	• Human assessors can set out and explain their thought processes leading to a decision; computers cannot set out the rationale for their decision, nor answer questions posed to them.
DMI7b	• Because of the number of variables involved, data mining algorithms

	become little more than *black boxes*.
DMI8	Unclear whether a data mining algorithm can determine what is sufficient to constitute *reasonable suspicion* to justify a secondary search.
DMI9	A data mining programme only identifies correlations between an individual a profile. It does not represent causation explaining why that individual possesses those characteristics or has performed the activities they have done.
DMI10	Data mining produces administrative decisions without due process; there is no evidence, no opportunity to be heard, no cross examination of evidence or opportunity to present counter-evidence - there is only a result.
DMI11 *DMI11a*	Data mining administrative decision lack the benefit of human judgement and human experience; • No ability for data mining algorithms to weigh evidence, apply knowledge from the real world, the current situation, and current events, beyond its programming so as to inform their decision.
DMI12	Algorithms have no brain and cannot interact, thus they will not change their decisions based on reasoned arguments or additional information.
DMI13 *DMI13a* *DMI13b* *DMI13c* *DMI13d* *DMI13e* *DMI13f*	Garbage-in-garbage-out; the accuracy of decisions by data mining algorithms is directly influenced by the accuracy of the data that is mined. It is affected by: • The structure and consistency of the analysed data. • Duplicate records. • Lack of data standards. • How often data is updated. • Human error. • Subtle differences between data sets.
DMI14 *DMI14a*	Data mining across computer systems depends on interoperability between those databases; • Systems are not necessarily designed to be accessible by outside organisations/agencies which can limit their data mining usability and accuracy.
DMI15	Predictive data mining processes are increasingly being used to make decisions about the criminality of individuals even though they possess no-where near the required level of accuracy, thus are seen as civil liberty infringing engines.

DMI16	Data mining is used for *fishing expeditions* where authorities *search* before they *suspect* which goes against our presumption of innocence.
DMI17	The negative repercussions of false positives will be repeatedly felt by those selected by the data mining algorithms; thus the individual is repeatedly being falsely accused or suspected.
DMI17a	• This is compounded when these algorithms are shared between government departments or with the private sector.
DMI18	Data mining algorithms are not ethically neutral; if the database data is biased then the results of the algorithm will also be biased.
DMI18a	• Decisions over what variables to include or exclude from the algorithm are human subjective judgements made by people with their own biases.
DMI19	Civil liberty arguments pertaining to data mining include the following;
DMI19a	• Data mining processes have a chilling effect on political expression as individuals become worried about repercussions/consequences from engaging in legitimate protests.
DMI19b	• Observation and surveillance by others limits the autonomy of individuals by repressing dissent.
DMI19c	• Weakening privacy controls allows for data to be collected from companies by the government which can then use that data for purposes other than for which it was collected.
DMI19d	• Data mining across different sources can make explicit that which was implicit, thus negating an individual's privacy and opening them to exploitation.
DMI19e	• Privacy protects people from being abused by others, including governments.
DMI19f	• Data mining alters the *status quo* between citizens and government; the government has more knowledge/power than the citizen.

2.9 Data Matching

Data-matching can be conceived as "*matching individuals with data about them*" (Steinbock 2005, p.4) which historically covered such activities as the manual comparison of the fingerprints of a suspect with those lifted from a crime scene, or a witness looking through books of *mug-shots* on the chance they recognise someone. While such activities still occur today, when I refer to data-matching I am restricting

this to computer-based systems which have automated the process. Examples here include comparing the identity of prospective airline passengers with those on *no-fly lists*, or the undertaking of DNA sample comparisons against those held in a national database. As such for the purposes of this examination I am employing the definition of data-matching adopted by Steinbock, that being "*the computerized comparison of two or more systems of records*" (2005, p.10). The obvious change brought about by the computerisation of data-matching is it greatly enhances that number of comparisons which can be made and reduces the time this takes.

Data-matching can be employed when addressing both *past-events* such as using biometric evidence to match an individual to a crime scene, and *future-events* such as matching a suspected terrorist to a flight watch list which triggers additional airport screening even though no offence may have been committed at this point.

2.9.1 Purported benefits of, and justifications for, data matching

Automated data-matching allows for the comparison of individual pairs or sets of data at a rate which was physically unachievable when conducted manually, regardless of resources. This matching speed makes the technology operationally feasible for busy setting such as airports and sporting events.

Automated data-matching technology has made possible biometric identifier systems; such as DNA databases and biometric passports. Non-biometric systems include automated number-plate recognition systems for both tracking movement and for identifying vehicles which are not insured or are stolen.

2.9.2 Associated problems with the design of data matching

While it is noted that data-matching is likely to produce more accurate results than data-mining due to the simpler algorithms involved, it is not without its problems. Matching data from different databases, each with their own system of data entry and storage, can result in real problems arising from the most innocuous of factors, such as the inclusion or exclusion of a second initial. This is before one even considers the fact that many people throughout the world will share their name with at least one other

person; a fact which becomes more pertinent when databases grow to include millions of people and links different countries. People have been subjected to secondary screening or even the denial of services because their name matches (or even resembles) aliases used by suspected terrorists (Steinbock 2005).

Another factor which undermines the reputation of data-matching procedures is when watch-lists become so large that they lose public credibility, calling into doubt the true purpose of the lists themselves and raising questions as to the underlying assumptions and/or quality of the data used to justify the inclusion of an individual on such lists. To illustrate, by 2005 the terrorist watch list for American airports (both 'no-fly' and 'selectee' lists) contained 70,000 names, while the US Terrorist Screening Center (TSC) database held over 230,000 names as potential terrorists (Steinbock 2005) out of an approximate population of 295 million[86]. This large number can be explained by the position of the TSC to include anybody on this list where any degree of terrorism nexus could be shown; hence the number of false positives would potentially be staggering.

Once an individual is nominated for inclusion on a watch-list by any number of sources or government agencies there is no further review by the database administrator; the individual is simply included. Nor is the information justifying the application for inclusion made available for integrity checking by the database administrator or made available at the point of screening to those who would be charged with undertaking this secondary screening to allow the application of human intuition (Steinbock 2005).

Given that watch-lists may be shared between different government departments and with private sector organisations, membership of a watch-list as a false positive can result in one experiencing multiple sanctions far beyond those applied by the original organisation (Kreimer 2004-2005).

2.9.3 Reactions and responses to data matching problems

As a *process*, data matching has not elicited the same level of social resistance as some of the other technologies discussed in this chapter, and hence there has not been the impetus for any overarching official remedial responses to date. Rather, negative

[86] Estimated 2005 US population according to www.census.gov/popclock/

public reactions have focussed on specific applications of this process; with the prime example being the UKs national DNA database (NDNAD).

The NDNAD contained 6,969,396 DNA profiles relating to an estimated 5,950,612 individuals as of 31 March 2012[87] and is continuing to grow in size. The scale of the NDNAD can be attributed to the legislative framework governing its operation[88] which allows the police "*to take and to retain indefinitely, without consent, fingerprints and DNA samples from a person of any age who has been arrested in connection with a 'recordable offence'*" (Hepple 2009, p.78). These samples remained even if the person was not charged with an offence, had the charges dropped, or if they were subsequently acquitted. Following the European Court of Human Rights (ECtHR) case of *S and Marper v The United Kingdom*[89] whereby the court unanimously held this system violated Article 8 ECHR the UK Government reformed the DNA collection and retention process via the Protection of Freedoms Act 2012. This will both restrict those samples which can be retained as well as setting time limits on many of the retentions.

2.9.4 Identified controversies arising from data matching

Table 2.8: Identified controversies arising from data matching

Code	Technology = **Data Matching**
DMA1	Data matching across increasingly large databases will result in more people becoming false positives as the total population size of the combined databases grows;
DMA1a	• When this results in secondary screening or denial of services, more and more people will be adversely affected.
DMA1b	• When databases are shared these effects will be felt by the individual across a range of services and situations.
DMA1c	• The negative repercussions of false positives under data matching systems will be repeatedly felt by those selected by the algorithm, thus the individual is repeatedly being falsely accused or suspected.
DMA1c1	○ This is compounded when these algorithms are shared between government departments or with the private sector

[87] As per House of Commons Hansard Written Answers for 19th June 2012: Column 866W.
[88] Namely the Criminal Justice and Police Act 2001.
[89] *S and Marper v United Kingdom* [2008] ECHR 1581.

DMA2	Watch-lists which become increasingly large lose public credibility and call into question the true purpose of the lists themselves and the quality of the data being used and assumptions made.
DMA3 *DMA3a*	When individuals are nominated for inclusion on watch-lists no further reviews or integrity checks by database administrators are carried out; • The information justifying inclusion is not made available to the secondary screeners to allow them to apply human intuition.
DMA4	Data matching produces administrative decisions without due process; there is no evidence, no opportunity to be heard, no cross examination of evidence or opportunity to present counter-evidence - there is only a result.
DMA5	Data matching processes have a chilling effect on political expression as individuals become worried about repercussions from engaging in legitimate protests against the prevailing government or policies.

2.10 Closed Circuit Television

Kroener (2010) describes closed-circuit television (CCTV) as "*the use of a video camera system to transmit a signal to a specific monitor or set of monitors (as opposed to a public broadcast source)*" (p.11). The first generations of CCTV systems utilised as STs possessed limited functionality, operating solely as monitoring system; either in *real-time* (whereby the end-user could look at a monitor connected to the CCTV camera to see what has happening at that moment in time) and/or *displaced in time* (whereby after the technology became available, CCTV systems incorporated a recording capability thereby allowing the recorded images to be viewed at the end-users discretion).

Today such cameras are used for a wide variety of security purposes by different end-users: "*home owners [install cameras] on their gates in order to see visitors (or intruders) on a screen inside their houses. Businesses use them for the protection of commercial premises, and shopping centres have cameras installed with images watched by staff in a control room*" (Kroener 2010, p.11). The state and its agents employ cameras in public and private spaces to deter, detect, and monitor the occurrence of offences; from simple traffic violations to counter terrorism operations.

However, moving beyond these initial iterations of CCTV systems, the latest generations have progressed past being simple monitoring/recording instruments. They can operate as platforms to enable the application of additional types of STs. This includes; biometric technologies[90], ATD systems, profiling systems, persistent authentication technologies, ANPR capabilities, amongst others (see Introna and Wood 2004; Burghouts et al 2011; McKenna and Gong 1997; and Thiel 2000 for expanded discussions on these applications).

2.10.1 Purported benefits of, and justifications for, closed circuit television

The benefits of CCTV are purported as the following (Gill 2003):

- They aid in police investigations;
- CCTV provide intelligence about an offender's actions both before and after an offence;
- Footage can encourage a guilty plea, especially when the defendant is clearly identifiable;
- They can help enforce traffic and parking regulations;
- They allow police to manage large crowds and sporting arenas;
- They help protect the vulnerable.

It is also claimed that CCTV reduces crime, makes communities feel safer, and is supported by around 90% of the UK population (Gill 2003; Mackay 2003).

2.10.2 Weaknesses and drawbacks of closed circuit television

Exaggerated claims of efficacy

Despite the money invested, and the broad range of positive effects attributed to CCTV, a 2005 Home Office evaluation acknowledged that CCTV had been oversold as the answer to problems of crime by successive governments (Gill and Spriggs 2005). This 'overselling' may be fostering a negative long-term environment for CCTV. As Gill

[90] Such as facial recognition systems.

notes, "*paralleling the widespread use of CCTV have been growing doubts about its capacity to deliver on early promises of effectiveness … it has yet to be proven that the benefits outweigh the drawbacks or that CCTV is cost-effective*" (2003, p.1). Mackay (2003) claims there is little evidence that CCTV in town-centres have achieved their promises of reducing crime or the fear of crime, nor have they increased public confidence. And in 2007 DCC Graeme Gerrard[91] admitted weaknesses with CCTV, specifically its failure to prevent crimes.

Even when CCTV succeeds in identifying a criminal act in progress this does not mean an appropriate response will necessarily ensure. Gill (2003) notes that awareness of an offence occurring and the ability to respond effectively are two separate matters, especially in rural areas where distance is a factor, in out-of-town industrial parks when police are focussing their efforts on inner-town areas (Gill and Loveday 2003), or where the police are fully stretched such that there are no police resources available to respond at all such as during the 2011 London riots.

Crime-reduction and crime-displacement claims

According to police sources, in 2008 only 3% of street robberies were solved through CCTV as, amongst other things, 80% of the produced images were of such poor quality as to be useless (Hempel and Topfer 2009). Additionally the evidence behind claims that town-centre CCTV systems are good crime prevention measures is often disputed. Conflicting evidence claims CCTV has a minimal effect, only working as part of a package of crime prevention measures.

Feelings of safety

Evidence would suggest that the presence of CCTV does not make people feel safer, undermining one of the central argument of CCTV proponents. This was the finding of studies in Glasgow and Hamburg (Ditton 2000; Mackay 2003; Zurawski 2010). While there may be strong public belief that cameras will reduce crime before they are introduced, in a study by Gill et al (2007) this belief dropped off significantly after introduction with respondents "*significantly less happy with the cameras than they had anticipated*" (p.319).

[91] Former Deputy Chief Constable of Cheshire.

The existing consensus appears to show no simple direct causation between the introduction of CCTV and the attitudes of citizens as to their feelings of safety. Rather safety in public places is recognised to be a complex issue incorporating complex social dynamics, fears, experiences, and attitudes, which cannot all be simply addressed by the introduction of CCTV (Gill et al 2007; Zurawski 2010).

CCTV as revenue raising tools

Gill (2003) notes that there is concern by the public that CCTV is being used to raise revenues [through the imposition of fines]. While public support exists for the crime-reduction/public-safety functions of CCTV, this does not extend to CCTV being used as a tool for raising money.

CCTV being used inappropriately or not complying with legal requirements

Within the UK, CCTV use is primarily governed by the Regulation of Investigatory Powers Act 2000 (RIPA) and the Data Protection Act 1998 (DPA).

RIPA was enacted to regulate the use of investigatory powers such as those used by police to tackle crime, and by extension terrorism. In doing so RIPA covered the use of covert surveillance by local authorities when fulfilling their statutory functions (such as preventing environmental crime, loan sharking, employment of minors, etc.). However local authorities also use these directed surveillance techniques when tackling dog fouling, littering, and for policing school catchment areas (BBC News 2008a, 2008b & 2009a) resulting in negative public and media reactions, and in RIPA being dubbed the *snooper's charter*.[92]

The DPA places legal requirements on the use of CCTV. In a study of CCTV operations by McCahill and Norris (2003) only 43 of 81 premises surveyed were displaying the legally required signage stating the purpose of the surveillance and contact details of the data controller. They estimate that 78% of CCTV systems in London businesses were not compliant with the DPA.

Apart from the DPA and RIPA there is very little to regulate the use of cameras in the UK prompting Norris and Armstrong (1999) to note that this proliferation of public

[92] The government's Protection of Freedoms Bill (2011) plans to ban councils using RIPA powers unless signed off by a magistrate and required for stopping serious crime.

space cameras has not been facilitated by the passing of legislation but by the absence of legislation regulating or checking their installation.

Proliferation of CCTV and associated liberty issues

Extrapolating from a study of CCTV cameras present in two streets in Putney, it was estimated that there are at least 4,285,000 CCTV in the UK (McCahill and Norris 2003). While this number has not been confirmed there is a concern that if citizens begin to doubt the effectiveness of CCTV as a crime prevention measure then they will see these cameras as a *big brother* measure for spying on citizens (Gill 2003).

Looking beyond the crime prevention *raison d'etre*, CCTV has been defined as a range of technologies aimed at classifying citizens, uncovering deviance, and inducing conformity (Norris and Armstrong 1999). CCTV systems are increasingly being used as platforms for, or facilitators of, additional technologies such as automatic number plate recognition (ANPR) and facial recognition which moves into the area of managing *risk*. Risk makes everyone a legitimate target for surveillance; "*[e]veryone is assumed guilty until the risk profile assumes otherwise*" (Norris and Armstrong 1999, p24). This provides a convenient avenue for addressing the question put forward by those against CCTV of 'why am I being recorded if I am not doing anything wrong?'.

A danger for civil liberties is the potential for misuse of CCTV. Street cameras have been used by council workers to spy inside houses (BBC News 2006), and authors have identified situations where operators discriminate by disproportionately monitoring the movements of young people or ethnic minorities (Loveday and Gill 2003).

Public support for CCTV is overinflated

A 1993 survey in Glasgow reported 90% support for CCTV in Glasgow city centre which has become a mantra for those who support such systems; however this survey lacked methodological rigour being conducted by post. Subsequent surveys placed the level of support at between 64% and 67%, nevertheless this 90% figure has become an iconic number (Mackay 2003; Norris and Armstrong 1999; Smith et al 2003). Indeed it is the view of Norris and Armstrong that it is impossible to ascertain the true level of support for CCTV.

Using these headline figures without further examination may also provide a false picture of the public's support for CCTV. As Zurawski (2010) notes, a study where 66.4% of respondents supported CCTV only 32.1% offered absolute support, implying that the remaining 34.3% were not fully supportive, had doubts or only offered conditional support. Only 50% of those respondents were in favour of constant camera operation, 16.3% supported night-time only monitoring, and 6.7% monitoring of rush hours. The point being that CCTV support varies considerably when the question is rephrased into more detailed and nuanced questions.

2.10.3 Reactions and responses to the closed-circuit television problems

Legislation (both domestic and EU-wide) governing the acceptable usage of CCTV has been the main response by governments to concerns over this technology. Such legislation may have bespoke application to CCTV or possess more generalised application. Within the UK these include; the Data Protection Act 1998, Regulation of Investigatory Powers Act 2000, the Protection of Freedoms Act 2012, and the Human Rights Act 1998.

The main European directive addressing CCTV is EU Data Protection Directive 95/46/EC. There are attempts underway for this to be superseded by a single Europe-wide law in the form of the General Data Protection Regulation though the possible adoption of this regulation is not planned until 2014.

To meet their obligations under Article 28 Directive 95/46/EC, as given effect within the UK by the Data Protection Act 1998, the Information Commissioner's Office was created. This office (amongst other things) provides an independent body to uphold the information rights of individuals, including those in relation to CCTV.

2.10.4 Identified controversies arising from closed circuit television

Table 2.9: Identified controversies arising from closed circuit television

Code	Technology = **Closed Circuit Television**
CTV1	Over-exaggerated claims of efficacy means the technology is not able to

	fulfil its promises resulting in a fall in public support.
CTV2 *CTV2a*	Live CCTV does not guarantee a timely response even when an operator identified a criminal act in progress; • Distance and lack of resources to respond will be factors here.
CTV3 *CTV3a* *CTV3b*	Poor image quality and/or poor positioning of CCTV cameras will negate the effectiveness of this technology; • Inability to identify the offender. • Lack of clarity will negate attempts to secure a conviction in court.
CTV4	CCTV in city centres should be part of a package of crime prevention measures to be effective; it is not a *silver bullet*.
CTV5	Evidence suggests CCTV does not make people feel safer once implemented.
CTV6	When CCTV is used for crime prevention/public safety goals it enjoys public support. When it is seen as a *tool for raising money* it does not enjoy public support.
CTV7	Covert CCTV used by councils for tackling dog fouling, littering, or to police school catchment areas does not enjoy the support of the public and is seen as a disproportional use of powers.
CTV8	CCTV street cameras have been used by council workers to spy on women in their homes.
CTV9	The majority of CCTV does not comply with the Data Protection Act 1998 requirements.
CTV10 *CTV10a* *CTV10b*	Estimated there are at least 4,285,000 CCTV cameras in the UK; leading to claims: • They are for spying on citizens. • The interests of businesses are being promoted over the rights of citizens.
CTV11 *CTV11a*	Creation of a *ring of steel* type CCTV and ANPR system around two semi-residential areas in Birmingham without public consultation resulted in significant community anger and loss of trust in the police; • Deliberately lying about the true purpose of a CCTV system (anti-terrorism) and hiding it behind a smokescreen of non-deliverable alternatives purposes (anti-crime, safer environments, etc.) destroyed public trust and support for the scheme.
CTV12	A system of covert and overt CCTV and ANPR cameras permanently ringing a predominantly Muslim semi-residential area and set up for the purpose of

CTV12a *CTV12b* *CTV12c*	surveying the residents as part of a counter terrorism measure represents a failure to balance the rights of privacy and security; • Police failed to consider the ethics of ringing a community with covert and overt CCTV and ANPR cameras. • They did not consider the effects this would have on the residents. • Those within the police who would have understood the consequences of this were not informed of the project.

2.11 Hand-Held Explosive Detectors

On numerous occasions since the 1990's devices have appeared in the security market claiming to be able to detect from a distance everything from explosives, drugs, weapons, money, and ivory through to humans lost at sea or trapped under rubble. Originally these devices appeared as the Quadro Tracker which was marketed and sold to US local and national law enforcement authorities for US$995 per unit as a detection device for weapons and drugs. Other reincarnations include the GT200, ADE-651, Mole detector and the SNIFFEX Handheld Explosives Detector.

Original image removed for copyright reasons from this electronic version.

Image can be viewed online at:

http://www.mirror.co.uk/news/uk-news/british-businessman-sold-golf-ball-1750300

Figure 2.2 The ADE-651 detector produced by ATSC Limited.

The physical construction of these devices is often similar. They are invariably hand-held, usually contain no batteries rather claiming to be powered by static electricity generated by the user, may possess a slot for inserting some form of *programmable substance detection card* depending on what the user wishes to find, and will all have some form of swivelling antenna pointing out at right angles to the user which moves in the users hands to point at whatever substance the device is programmed to search

for. In effect these devices are the modern day equivalent to dowsing rods; an ancient superstition whereby people used forked sticks to look for underground water sources.

These hand-held detection devices also have other similarities. Firstly they are very expensive, selling for as much as £45,000 each. The Iraqi government spent £52m purchasing ADE-651's from ATSC Limited, a UK company based in Somerset (Hawley and Jones 2011a; Sengupta 2010). Secondly they are marketed as being hard to use and requiring extensive operator training; often due to the devices' supposed sensitivity, as well as the requirements placed on the user that they be relaxed and/or in a certain mental state before the device will work.

These devices are distinct from the explosive trace detection systems often found in airports, and hand-held explosive *sniffer* devices which employ a range of technologies and scientific principles to enable detection. This includes; ion mobility spectrometry, thermo redox, chemiluminescence, and amplifying fluorescent polymers (Ghosh et al 2010; Thomas III et al 2005). The sensitivity, cost, and portability of these devices varies.

2.11.1 Purported benefits of, and justifications for, hand-held explosive detectors

According to the marketing literature produced by those selling these products[93], hand-held explosive detectors possess the following attributes.

1. They can detect any known drug or explosive substance for which they are programmed to detect.
2. They employ electrostatic ion attraction technology to target these substances; i.e., the electro-static matching of the ionic charge to the structure of the substance.
3. They require no internal power source, rather are powered by the static electricity generated naturally by the user.
4. They ignore all known concealment measures, detecting substances through lead, other metals, concrete, and within the human body.

[93] See Appendix *D*.

2.11.2 Associated problems with the design of hand-held explosive detectors

Hand-held explosive detectors of the type described here have never been shown to work at detecting whatever they claimed they could detect in any independent trial, ever. Double-blind scientific experiments have all shown that they perform no better than random chance at detecting specific materials (National Institute of Justice 1999). Tests on the SNIFFEX by the US Naval Explosive Ordinance Disposal Technology Division (2005) resulted in the conclusion that "*the SNIFFEX handheld explosives detector is not capable of detecting explosives regardless of the distance between the device and any explosives….The SNIFFEX handheld explosives detector does not work*" (pp.7-8). These facts have resulted in arrests and prosecutions of their manufacturers for fraud and bans on their sales.

In short there devices do not work. They represent a scam which in the past has taken in many countries including the United States, Iraq, and Thailand. This fact could be used as justification for not including these detectors as a case-study here as these devices are *not actually a technology*[94]. Nevertheless, these devices are often treated as a legitimate ST by both governments and end-users. They continue to take in new countries and even today they remain in operation in Mexico, Kenya, Lebanon, Jordan, and China amongst others (Hawley and Jones 2011b). While this 'scam' claim could be made about many products available on the internet, it is the life threatening repercussions arising from the failure of this technology that sets it aside from the others. In reporting on the true nature of the ADE-651 which is used in dozens of Iraqi checkpoints in cities including Baghdad, Sengupta noted that "*[i]t is claimed that it failed to detect two tonnes of explosives used by suicide bombers to murder 155 people and destroy three ministries in October [2009] …. There was a similar alleged shortcoming when 120 people were killed in another series of bombings*" (2010). It is for these reasons, and the controversial nature of these devices, that I justify including them as a case study.

[94] This claim is made using the definition of 'technology' is defined in the Oxford Dictionary as the application of scientific knowledge for practical purposes

The anger which follows such tragedies upon discovering the truth of security scams may be directed far beyond the devices themselves and is something which also must be accounted for. An Iraqi teacher whose son was killed in one the bombings described above said:

> I am angry. I do not know who I am angry with more, the people who made these stupid things and then made money or our government officials who paid so much money for these things which failed to protect us. And the British Government, did they not know what was being done from their land? (Sengupta 2010).

2.11.3 Reactions and responses to hand-held explosive detector problems

Both the US and UK governments have publically acknowledged the ineffectual nature of these particular devices.

The US Federal Bureau of Investigation (FBI) along with Sandia national laboratories identified the Quadro Tracker as a fraud in 1995 warning all US agencies immediately cease their use; a warning they repeated in a 1999 report. Following this in 2002 the Mole detector, a reincarnation of the Quadro, was also identified as a fraud by Sandia national laboratories (Hawley and Jones 2011a).

Within the UK the ADE-651 built by ATSC which has been sold to Iraqi government was the subject of two BBC Newsnight investigations, eventually resulting in the police arresting the company's managing director James McCormick on suspicion of fraud by misrepresentation, and the Department for Business, Innovation and Skills banning the export of this device to Iraq and Afghanistan. This is the limit of their legal powers, hence the limited nature of the bans (Sengupta 2010).

McCormick's trial commenced on the 6th March 2013 where he was charged with three counts of fraud for selling bomb-detectors which he knew did not work (Daily Mail 2013b). He was subsequently found guilty and sentenced to 10 years in jail (BBC News 2013; Booth 2013).

2.11.4 Identified controversies arising from hand-held explosive detectors

Table 2.10: Identified controversies arising from hand-held explosive detectors

Code	Technology = **Hand-Held Explosive Detectors**
HED1	These devices are a scam and do not work. Independent tests have always shown them to be a scam.
HED1a *HED1b*	• The damage is compounded by the fact that people believe they are security enhancing devices, providing protection by detecting explosives. • Hundreds of people have died in Iraq from insurgent bombings in buildings/areas supposedly protected by these fake devices.
HED2	People rely on these devices and make decisions regarding personal risk based on the false belief that they do work.

2.12 Mosquitos

Mosquitos are electronic auditory security devices which emit a pulsed high frequency tone, the unpleasant nature of which is such that those within its range are commonly forced to disperse after a short period of time. It has two frequency settings; either around 17.6KHz which can usually only be heard by those under 25 years of age with a range of up to 40m, or 8KHz which can be heard by everybody regardless of age over a longer range of up to 60m. The reason those over 25 cannot usually detect the higher frequency tone is *presbycusis*; the naturally occurring phenomenon whereby humans lose the ability to hear high frequency sounds as they get older.

2.12.1 Purported benefits of, and justifications for, Mosquitos

The Mosquito is specifically marketed as a method of reducing teenage graffiti, vandalism, loitering and anti-social behaviour. This is based on the assumption that as this device makes the immediate area an unpleasant and annoying place for teenagers to be, they will leave the area and these activities will cease.

According to the manufacturer 'Compound Security' the Mosquito MK4 operating on the higher 17.6KHz frequency setting is:

> the most effective tool for dispersing groups of teenagers who loiter and behave in an antisocial manner… without confrontation! If you have problems with teenagers loitering near your property, causing criminal damage, putting off customers or abusing your customers and staff, the Mosquito MK4 is the most effective method of putting a stop to it (Compound Security (a), no date).

The 8KHz frequency setting can be heard by everyone and is designed for use in areas such as subway tunnels, car parks, etcetera, where "*homeless people sleep rough … [and] unpleasant characters gather at night*" (Compound Security (a) & (b)). Compound Security go on to assert that Mosquitos pose no health concerns, even with prolonged exposure, and are perfectly legal to both own and use (Compound Security (a)).

Their use is supported by the Association of Convenience Stores[95], some retailers and police forces (Crawford 2009). The Labour government in 2010 indicated support for these devices, rejecting calls for a ban saying it is up to local Councils and police forces to regulate their usage. Similarly to date there has been no national legislation introduced regulating how Mosquitos are to be used. The end result being over 5000 units have been sold in the UK (Peck 2010), with no figures available for international sales. Even critics of Mosquitos acknowledge their apparent effectiveness based on the testimonials of satisfied Compound Security customers (Walsh 2008).

2.12.2 Associated problems with the design of Mosquitos

While around 25% of UK local councils and police forces admit to using or endorsing the Mosquito, support is far from universal. The Association of Chief Police Officers (ACPO) refuses to nationally approve the device citing safety concerns, while the Council of Europe, human rights group Liberty International, former England children's commissioner Sir Al Aynsley-Green, the Children's Society and the Children's Rights Alliance for England, amongst others, all oppose these devices on various legal, ethical, moral and social grounds (Peck 2010).

[95] Caroline Gall, BBC NEWS – 30 June 2010 – *Stafford teenager fighting for mosquito device ban*. Last accessed 12/05/2011 from http://www.bbc.co.uk/news/10449634

The question of legality

Within the UK while there is no national law specific to Mosquitos either restricting or regulating their use; conversely there is also no law expressly permitting their use either. Despite this certain local authorities have banned Mosquitos, including Kent, Edinburgh[96] and Stirling[97].

Aside from these localised restrictions, the legality of Mosquitos may be susceptible to legal challenge through existing non-specific legislation. According to Walsh (2008) these include:

- s.79 Environmental Protection Act 1990 which requires local authorities inspect their area for *noise nuisances*,
- s.1 Crime and Disorder Act 1998 allowing for the imposition of an Anti-Social Behaviour Order (ASBO) against the operator,
- s.1 Protection from Harassment Act 1997 whereby '*a person must nor pursue a course of conduct which amounts to harassment of another; and which he knows or ought to know amounts to harassment of the other*',
- and the potential breach of both the common law tort and crime of *public nuisance*.

Additionally there are arguments against Mosquitos under the Human Rights Act 1998 (HRA) which incorporated most of the European Convention on Human Rights (ECHR) into UK law. While the HRA applies to *public* not *private* bodies (so not shopkeepers), public bodies may become liable for failing to protect individuals against the actions of other individuals. The HRA rights most likely to be engaged according to Walsh (2008) are:

- Article 8 *right to respect for private and family life*,
- Article 11 *freedom of assembly and association*, and
- Article 14 *prohibition of discrimination*.

[96] 'Is the Mosquito Alarm an Infringement on Human Rights?'. From Civil Rights Movement website, last accessed online 03/05/2010 from http://www.civilrightsmovement.co.uk/mosquito-alarm-infringement-human -rights.html
[97] See BBC News (2008c).

Moving away from specific arguments based on UK legislation, questions have also been raised about the status of Mosquitos contravening various international Conventions of which the UK is a signatory. The Council of Europe (CoE), voted on the 25th June 2010 to ban the marketing, selling and use of Mosquitos in all public places throughout CoE member states on the basis that such devices[98]:

- Are illegal solutions under international human rights instruments,
- Demonise young people,
- Cause young people to lack confidence in the legal system,
- Constitutes a possible health hazard,
- Does not solve the problems facing young people.

While the CoE does not have law making powers the views of the CoE will be persuasive should the ECtHR be asked to consider the position of Mosquitos in relation to the ECHR. Finally the United Nations Committee on the Rights of the Child (which oversees the operation of the United Nations Convention on the Rights of the Child (UNCRC) of which the UK is a signatory) raised concerns regarding Mosquitos violating Article 15 UNCRC: freedom of association.[99]

Moral and social issues surrounding the use of Mosquitos

The design and operation of Mosquitos raises several moral and social questions of much wider scope than simple legality.

1. *What does this technology say about how UK society views its children?*

Crawford, in a single paragraph provides a succinct damning view on this question:

> The device purports to afford a technological means of dispersing youths regardless of their motivation or behaviour in an impersonal and indiscriminate way. It does so without any notion of what to say to them, how to engage and reason with them or even how to socialize them. It lacks any attempt to inculcate pro-social behaviour or moral values, but instead emits a droning noise that

[98] Council of Europe, Doc.12186 'Prohibiting the marketing and use of the "Mosquito" youth dispersal device', 22 March 2010. Accessed online at http://assembly.coe.int/Documents/WorkingDocs/Doc10/eDOC12186.htm on 05/05/2011.

[99] UN Committee on the Rights of the Child, *Consideration of reports submitted by states parties under article 44 of the Convention*. Document CRC/C/GBR/CO/4, 20 October 2008, paragraph 35.

> implicitly says 'go away'. This would appear to reflect a rather hollow approach to young people on the part of adult society. (2009, p.21)

The *Buzz-Off* campaign, a joint initiative of Liberty, 11 Million and the National Youth Agency, was formed to have Mosquitos banned. This campaign argues that one of the reasons why Mosquitos should be banned is that they demonize young people, creating divisions instead of mutual understanding and respect.

2. *Who owns public spaces?*

The purpose of the original Mosquitos (those with only the higher frequency setting) was to disperse teenagers, usually from public spaces. Those against Mosquitos argue that young people have as much right to be in public spaces as everybody else.

By distilling the advertising literature and arguments of those promoting Mosquitos the results are predominantly economic arguments, whereby the presence of teenagers is viewed as either potentially impacting upon commercial activities or as a source of past and future damage. But this direction of argument either diminishes or ignores two important factors. Firstly, if teenagers are not engaged in illegal activities then what right do shopkeepers have of forcing them out of public spaces? As Walsh comments, "*it should not be the role of shopkeepers to determine who can legitimately use public space*" (2008, p.132). Secondly, where exactly are they expected to go? Crawford (2009) illuminates the inherent paradox for young people in dispersing them from public places in that they often congregated there to feel safe; a safety they derived from the feeling of gathering together and of being in public spaces, visible to others. By forcing them to break up and disperse, these youths felt at greater risk as they were forced into locations with less adult foot-traffic and were often less well lit and less safe.

3. Equality of burden

By developing the 17.6KHz Mosquito detectable only by the young, the manufacturers have succeeded in intentionally creating a discriminatory security measure which targets a minority group. This form of discriminatory security technology reduces the liberties of a minority to provide additional security for the majority, and those suffering from reduced rights and liberties do not get to enjoy this additional security benefit.

The Compound Security Mosquito MK4 with Multi-Age has the option of emitting an 8KHz tone which can be heard by people of all ages. This device counters the first *equality of burden* issue identified in the preceding discussion in that ostensibly everyone will be equally affected by this technology. However on a deeper examination holes appear within this argument.

Firstly the 8KHz operation suffers the same problem as the 17.6KHz in that it does not discriminate between those engaged in illicit activities and those simply legitimately exercising their right to occupy public spaces or even walk along a public street. Secondly the advertising literature proudly boasts that the Mosquito MK4 is a particularly effective *solution* for targeting the homeless and those 'sleeping-rough' (Compound Security (a), no date). It is questionable whether a device which seeks to force such vulnerable members of society out of the sight (and mind) of the more fortunate in society represents a socially acceptable pursuit.

Operational limitations of the Mosquito

As Mosquitos are an *area-effect* device they cannot distinguish their effect between those engaged in illegal activities and those who are not. Secondly Mosquitos do nothing to solve the underlying causes of anti-social behaviour but simply displace it to different areas. Thirdly, when operating at the 17.6KHz frequency the Mosquito will do nothing to stop those too old to hear the tone from engaging in anti-social activities.

Lastly, not everybody can simply leave the area affected by the Mosquito. Babies react adversely to the tone but cannot communicate this fact to their accompanying parents who may themselves not be able to hear the tone nor realise what the Mosquito is or how it is adversely affecting their child. Young staff working within premises using the Mosquito may also be affected by the tone and again cannot simply walk away. Those with autism have also been identified as particularly susceptible to the Mosquito tone and may experience a greater adverse reaction than other members of the public.

Safety Issues

Mosquitos can cause dizziness, headaches, nausea, pain and impairment according to both the German Federal Institute for Occupational Safety and those subjected to them (Schreuder 2009). And yet determining whether or not Mosquitos are

completely safe in respect of exposure to children is not a question which can be answered with certainty due to the absence of specific data and the subjective nature of the effects induced by these devices.

Compound Security boldly claims the Mosquito is "*completely harmless to the health/hearing of individuals of all ages, even with prolonged exposure*" (Compound Security (a), no date), citing studies by the Health and Safety Executive, National Physical Laboratory and the Applied Environmental Research Centre as all confirming current legislation regarding the emission of high frequency sounds are complied with. Nevertheless closer examination of the sources behind these claims does not appear to justify the robustness of Compound Security's claims, especially in relation to the exposure of children.

Lawton (2001) produced a report on the damage to hearing from very high frequency (VHF) sounds for the Health and Safety Executive. Here he concluded there was insufficient data to produce a dose/response relation between risk of hearing loss and VHF sounds[100]. But he did recommend against any relaxing of current restrictions acknowledging that:

> In particularly sensitive individuals, unpleasant subjective effects might be expected to appear shortly after the start of a VHF or ultrasonic noise exposure. An increase of permitted band level … may be expected to hasten the onset of adverse subjective effects … and possibly … involve a larger proportion of the exposed population. (2001, p.46)

The lack of data pertaining specifically to children was highlighted throughout, as was the recognition that this group suffered greater to exposure to VFH sounds because of their heightened hearing ability. This lack of clarity regarding safety has been cited as the reason why ACPO refuses to endorse the Mosquito (Peck 2010).

2.12.3 Reactions and responses to Mosquito problems

Even though Mosquitos have not been subject to legal challenge within the UK, various local councils have restricted their use, including Stirling, Kent and Lancashire County Council. What is particularly telling here is that in Stirling Council the Mosquito was

[100] VHF sounds cover those between 10 and 20KHz, this incorporate the 'teenager focussed' Mosquito frequency band.

banned after complaints by a 26 year old Councillor who could still hear the noise emitted (BBC News 2008c). This brings into question the long-term viability of security measures designed specifically to target the young. As the first generation of previous 'targets' get older and enter positions of power they make take a dim view of such measures which they found degrading in their youth and take action against them.

First Great Western stopped using a Mosquito during the normal operating hours of a Devon railway station following complaints by school children who used the station on their travels to and from school. They admitted that legitimate travellers which included these teenagers should not have been subjected to the noise (BBC News 2009b).

Furthermore, according to the Civil Rights Movement some retailers have removed Mosquitos from certain stores once they were made aware of the device's harmful effects[101].

Nationally there have been no government plans to either ban or regulate the use of Mosquitos. This coalition position continues that of the previous Labour government.

2.12.4 Identified controversies arising from mosquitos

Table 2.11: Identified controversies arising from mosquitos

Code	Technology = **Mosquitos**
MS1	Mosquitos have been banned in certain local authorities.
MS2	Mosquitos may constitute a breach of s.79(1)(g) Environmental Protection Act 1990 if they constitute *noise emitted from premises so as to be … a nuisance*.
MS3	Operating a mosquito may constitute anti-social behaviour under s.1 Crime and Disorder Act 1998 thus permitting the imposition of an ASBO.
MS4	Mosquitos may constitute a breach of s.1 Protection from Harassment Act 1997 whereby *a person must not pursue a course of conduct which amounts to harassment of another*[102].

[101] See http://www.civilrightsmovement.co.uk/mosquito-alarm-infringement-human-rights.html last accessed only 05/05/2011.

MS5 *MS5a* *MS5b*	Under the Human Rights Act 1998 there are a number of potentially engaged rights: • Art.11 *freedom of assembly and association* – mosquitos prevent young people associating together in public places; • Art.14 *prohibition of discrimination* – mosquitos targeted at young people effectively discriminate against them on the basis of their age.
MS6 *MS6a* *MS6b* *MS6c* *MS6d*	The Council of Europe voted in 2010 to ban the marketing, selling, and use of mosquitos in all public places on the basis that such devices; • Are illegal solutions under international human rights instruments. • Can demoralise and frustrate young people. • They are a possible health hazard, targeting children and young people. • Do not solve problems facing young people, and are only negative towards them.
MS7	The United Nations Committee on the Rights of the Child considers mosquitos violate the rights of children to enjoy freedom of movement and peaceful assembly which are essential for the development of children.
MS8	Mosquitos reinforce negative stereotypes of children and adolescents, treating them all problems to be dispersed rather than human beings to be reasoned with.
MS9 *MS9a*	Mosquitos do not discriminate between youths who are and are not acting in anti-social ways; all youths are affected equally. • Law abiding children are affected in the same way as those committing offences.
MS10	Mosquitos alienate young people from their communities which could be counter-productive.
MS11	Mosquitos are degrading and the effect of the noise can be extremely uncomfortable.
MS12 *MS12a* *MS12b*	This technology raises questions about who *owns* public spaces; • If teenagers are not engaged in illegal activities then what right do shopkeepers have to force them away from public spaces? • It is not the right of shopkeepers to determine who can and can't use public spaces, especially as they possess a vested (economic) interest

[102] The s.1(3) defence of preventing or detecting crime may not apply here if the argument that 'these devices target teenagers rather than crime' holds sway.

	when making this determination so are not unbiased arbiters
MS13 *MS13a*	Youths often congregate in public spaces out of feelings of safety; • They may be placed at greater risk of harm if forced away from areas with high levels of foot traffic into areas which may be less well lit, frequented, and safe.
MS14 *MS14a*	Mosquitos only detectable by the young discriminate by intentionally targeting a minority group; • It imposes a greater burden on the rights of this minority group for the benefit of the majority.
MS15 *MS15a* *MS15b*	Multi-age mosquitos are disturbingly marketed as a *solution* to the problem of homeless adults who are *sleeping rough*; • A device designed to force vulnerable members of society out of the sight (and mind) of the more fortunate is morally bereft • It fails to tackle the underlying causes of why people are forced to sleep on the street.
MS16 *MS16a*	As mosquitos are an area affect weapon they cannot discriminate between those engaged in illegal activities and those engaged in legitimate activities; • This includes young shop staff who may feel forced to endure the mosquito out of fear of losing their jobs if they complain.
MS17	Mosquitos at the 17.6KHz frequency do nothing to stop the anti-social activities of those too old to hear this frequency of sound.
MS18 *MS18a* *MS18b* *MS18c*	Not all people who can hear the mosquito can simply leave the area, nor even communicate what the cause of their distress is. • Babies will be affected while their parents may not even be able to hear the sound that is distressing them. • Mentally disabled people may be affected to a greater extent without being able to verbalise the cause. • Mosquitos have been used in train stations affecting children going to school.
MS19 *MS19a*	Young people who were affected by mosquitos may be less likely to support their continued existence as they become older and enter into positions of authority; • A young councillor in Stirling Council who could still hear the noise was instrumental in having them banned by his Council.
MS20	Health concerns; mosquitos can cause dizziness, headaches, nausea, pain, and impairment.

2.13 Less Lethal Weapons (LLWs)

There is no single accepted definition of what constitutes a LLW. The North Atlantic Treaty Organisation (NATO) defines LLWs as:

> weapons which are explicitly designed and developed to incapacitate or repel personnel, with a low probability of fatality or permanent injury, or to disable equipment, with minimal undesired damage or impact on the environment (2006, p. 6-5).

The UK's Association of Chief Police Officers (ACPO) defines LLWs as:

> weapons, devices or tactics designed and intended to induce compliance without substantial risk of serious or permanent injury or death. The aim will be to control and neutralise a threat without recourse to lethal force. The outcome may occasionally be lethal but this is less likely than the result of the use of firearms (NATO, 2006, pp. 6-4 - 6-5).

As a general term LLWs covers a diverse range of technologies with varying effects (such as length and nature of incapacitation), delivery systems, effectiveness ranges, and effective targets (i.e., individuals, crowds or physical areas). Lewer and Davison (2005) have categorised LLWs by way of the technology which underlies how they are effective against their chosen target. These categories are; kinetic energy, barriers and entanglements, electrical, acoustic, directed energy, chemical, chemical/biological, and combined.

For any LLW to be considered effective it must be able to incapacitate, debilitate, or disrupt/alter the thought processes of either individuals or crowds. These effects must persist long enough so as to either prevent an advance or some other action, to cause them to voluntarily disperse, or to allow the relevant authorities to either restrain or disperse them (Council for Science and Society 1978).

As a final point, while a goal of LLWs is repeatedly cited as to raise the threshold for the application of deadly force, they do not prevent police or other security personal from using such force. Nor are police required to substitute LLWs for lethal weapons in situations where the use of lethal weapons would be legitimate.

2.13.1 Purported benefits of, and justifications for, less-lethal weapons

Purported benefits or various LLWs include the following:

- LLWs have proven particularly effective in controlling riots, including riots in prisons (Lewer and Davison 2005).
- Tasers could be cost-neutral by reducing the number of injuries to police and citizens, and their resulting medical and civil litigation expenses (Adams and Jennison 2007; Barfield 2010).
- Tasers operate as a legitimate alternative to lethal force (Sprague 2007).
- It has been identified that various LLWs have potential for use on aircraft to prevent 9/11 style attacks (Davison 2007) though this avenue has largely been ignored in favour of security advances elsewhere.
- There is the general argument that even though they can cause injuries, using kinetic projectiles such as rubber bullets is still better than using *real* bullets;
- There is evidence that the use of pepper spray and electronic control devices (ECDs) such as Tasers has reduced the number of minor injuries that were previously suffered from soft empty-hand controls when subduing actively aggressive subjects (Adams and Jennison 2007).
- LLWs save lives (Davison 2009) and improve society providing preconditions for their use are met (Braidwood 2009).

2.13.2 Weaknesses and drawbacks of less-lethal weapons

However there is also data indicating that while ECDs decrease the arresting officer injury rates they do result in relatively high suspect injury rates (Lin and Jones 2010).

1. LLWs as tools for suppressing political dissent and controlling populations

It has been argued that "*emerging non-lethal technologies offer an increasing opportunity for the suppression of civil dissent and control of populations*" (Lewer and Davison 2005, p.40). If citizens believe the use of these weapons undermines democracy (an example being the use of water cannons against civil rights protestors in the US during the 1960s) this may have a lasting and negative effect. As Downs (2007) notes, "*a water cannon may never [again] be acceptable in the USA as a less*

lethal weapon" (p.362). A second argument is that existing LLWs such as Tasers are used too readily by police as tools of compliance (Sprague 2007).

2. Threats to international arms conventions and international law

The Chemical Weapons Convention and Biological Weapons Convention prohibit the development of chemical and biological weapons respectively. However there is a caveat allowing for the research and use of chemicals for law enforcement purposes (Davison 2009). This has drawn considerable criticism by those who see the use of such chemical LLWs as potentially undermining the treaties themselves (Casey-Maslen 2010; Davison 2009; Lewer and Davison 2005). It has also been noted by these authors that no bespoke international agreements exist to restrict or regulate the development of acoustic and directed energy weapons.

3. Health concerns

Probably the most controversies arising from LLWs are the health concerns their use entails:

- *LLWs kill*: Despite the commonly employed moniker *non-lethal weapon* and the best intentions and efforts of the weapon designers the fact is LLWs can and do kill. There is no way to manufacturers can prevent this given they cannot control; (a) how their weapons are used in the field, (b) the physical characteristics of the target, or (c) simple bad luck. LLWs kill in two ways; (1) the process by which force is applied from the LLW to the target (i.e. kinetic, chemical, electrical, etc.) has the extreme reaction of killing the target, regardless of any *rules of use* for that weapon, and (2) people use LLWs to facilitate the killing of the targets. To expand on the first point, blunt impact weapons, Tasers, OC spray and calmatives have all killed the intended targets despite the best efforts of police forces and manufacturers to produce and enforce rules for their use, such as not firing Tasers into a subjects chest, or requiring medical assistance be called when OC spray is deployed (Braidwood 2009; Downs 2007; Sprague 2007; Stanbrook 2008; Taylor et al 2011; Welch 2011). On the second point, LLWs like CS have been used on battlefields to flush combatants out of strongholds so they can then be subsequently shot. Also in the siege of the Dubrovka Theatre in Moscow by

Chechen rebels, Russian security forces pumped an unknown aerosolised calmative agent (probably a fentanyl derivative) into the theatre incapacitating all the terrorists and the 800 hostages. This resulted in the deaths of approximately 130 hostages but also allowed the Russian security forces to summarily execute all of the 50 Chechen rebels while they were incapacitated (Davison 2009; Fidler 2005).

- *LLWs injure targets*: Any application of energy to a human body, be it kinetic, electrical, electromagnetic or chemical, is potentially harmful (NATO 2006). Evidence from the analysis of injury data also shows that "*suspects in ECD-deployed incidents [i.e. Tasers] had a higher injury rate than those in non-ECD incidents*" (Lin and Jones 2010, p.171).
- *The Lack of proper evaluations and the construction of biased evaluations*: When there are operational needs, pressure exists to get LLWs into the market and as a result a thorough evaluation is not undertaken (Lewer and Davison 2005). There is also the building controversy surrounding manufacturers of LLWs funding safety studies of their products where the independence of those conducting the studies is questionable. A knock-on effect of research which lacks perceived impartiality and independence is that people are now questioning the accuracy of reported device benefits (Lin and Jones 2010).
- *The variability of safety between devices*: Individual Tasers have displayed much higher output then should be produced when tested independently (Davison 2009).
- *The denial of health concerns*: The denial of health concerns has included the indefensible statement that no clear causal connection exists between impact projectiles and reported fatalities (see Vilke and Chan 2007) which is akin to denying a causal connection between being shot and dying from being shot. Perhaps the most vigorous denier has been Taser. Taser has; (a) consistently refused to accept their weapon affects the heart maintaining their weapons have never caused a death or serious injury, (b) sued a researcher for publishing peer-reviewed critical scientific evidence and a medical examiner for listing the Taser as a cause of death on a death certificate, and (c) created a refuted mental disorder (*excited delirium*) as an alternative to those deaths where a Taser was deployed

(Smith et al 2007; Welch 2011). According to Smith et al, Taser "*has demonstrated a willingness to squelch any message that could hurt its bottom line*" (2007, p.1402).

- *Physiological variability between targets*: It has been claimed that "*[i]ndividual differences [between targets] are perhaps the biggest single challenge for the designers of less lethal weapons*" (Downs 2007, p.377). Certain categories of individuals are vulnerable to a higher risk of injury and death from the various LLW, including amongst others, those who have taken drugs, pregnant women, the elderly, children, the mentally disturbed, asthmatics, and those with underlying cardiac conditions (Downs 2007; Fish and Geddes 2001; Sprague 2007).
- *Medical ethics concerns*: NATO (2006) acknowledges the inherent conflict for medical staff between the need to study the effects of prototype LLWs on human volunteers and the risk of death or causing harm these studies pose to the volunteers.

4. Function creep

Much of the controversy arising from the purported function creep of LLWs is borne from uncertainty over the true purpose for introducing these weapons in the first place; i.e., are LLWs alternatives to (or means of raising the bar for) the use of lethal force, or are they a means of forcing civilian compliance? As *forcing compliance* allows broader usage than restrictive situations such as those involving lethal force, allowing such usage will be interpreted as function creep if the LLW was *sold* to the public as an alternative to lethal force or for use in situations where the risk of serious harm is high. Using an example from Canada, the Braidwood Commission noted:

In West Australia the shadow attorney-general John Quigley was critical of Tasers and how the police were using these weapons. He was reported as saying "*[t]hey're (the police) using the Taser as a weapon of punishment. They're using the Taser as a weapon of control. They're using it as a weapon of compliance. It was never so intended*" (Marks 2010). According to Davison, "*enduring ethical and moral concerns remain over the use of weapons solely designed to cause pain at the push of a button*" (2009, p.182). These concerns are acutely felt by those doctors whose research into nociceptors (nerve cells that convey pain) with the aim of reducing chronic pain in

sufferers has been employed to help create weapons for hurting people (Editorial staff 2005).

5. Unacceptable usage and unacceptable targets

Electro-shock stun weapons including Tasers have enjoyed widespread use as tools for punishment and torture. Devices which attach electro-shock stun devices to the human body, such as *stun-belts*, are opposed by Amnesty International and the UK government as tools of torture (Fish and Geddes 2001; Smith et al 2007; Sprague 2007).

On the unacceptable use of Tasers, Amnesty International adopts the position that:

> unarmed suspects should not be shot with a Taser for arguing or talking back, being discourteous, refusing to obey an order, resisting arrest or fleeing a minor crime scene, unless they pose an immediate threat of death or very serious injury that cannot be controlled through less extreme measures" (Sprague 2007, pp.310-1).

This is in response to Tasers being used for just these situations. There is also the issue of the overuse of LLWs on certain ethnic groups. In Australia the use of Tasers against indigenous people between 2007 and 2010 was double that of non-indigenous targets (Guest 2010) raising questions over police racism. These questions were given a focal point when footage surfaced of an aboriginal man suffering from mental illness being shot 13 times while in custody and surrounded by 9 officers after refusing a strip search (Marks 2010).

6. Use of LLWs when alternative methods should have been deployed

The position that LLWs are more humanitarian than the alternatives is often undermined by discrepancies between how they are *intended to be used* and how they are *actually used*. They can subvert traditional policing practices by discouraging non-violent methods of persuasion in situations where lethal force would be unacceptable (Casey-Maslen 2010). LLWs can also increase the instances police *use force* and decrease the threshold for lethal force by using pain-causing LLWs to produce compliance in situations where conventional (physical) force would not be justified (Davison 2009).

Others have noted that as different LLWs have inherently different levels of injury risk the type of LLW used should reflect the nature of threat faced. If a target suffer serious injury from a LLW (such as a rubber bullet), then any argument that it was *better than a real bullet* will be undermined if the use of a *real bullet* would have been inappropriate in the given circumstances (Downs 2007).

7. The undermining of public support for LLWs

Public support is essential for the survival of any LLW. When unpalatable facts about how a particular LLW works and its potential side-effects come to light these can act to undermine the credibility and acceptability of that weapon. As NATO (2006) points out, "*[u]nacceptable facts or publicity can affect the public, politicians and the military user ... Policies will ultimately have to account for all elements of public awareness and acceptability*" (section 6-11).

Furthermore, incidents of user misuse may be highly publicised adding further fuel to the debate over these weapons. Without monitored and enforced policies for the use of LLWs to minimise such events, the resultant "*negative public dialogue could potentially detract from the many benefits of the device … and ultimately culminate in a moratorium on the use of the electronic control device altogether*" (Lin and Jones 2010, p.172). This is a fact now recognised by police forces that are becoming acutely aware of the levels of scrutiny under which they now operate.

Even when tightly drawn *rules for use* of a LLW permit its deployment, this may not be enough to convince a questioning public. As Smith et al (2007) have noted:

> no[] matter how unruly [the target, police] departments must consider how Tasering a minor or senior citizen will be viewed and digested by a highly critical media and public. … Tasering a pregnant woman or an individual in a wheelchair or on oxygen will not play well in the press or in front of a jury if the case makes it that far, even if an officer could otherwise articulate legitimate reasons for using the Taser under such circumstances" (p.410).

9. Lack of clearly defined rules for the use and monitoring of LLWs

There exists a lack of operational consistency over; (1) how Tasers are used, (2) where they fall of the police *use-of-force continuum*, (3) the level of training required, and (4) the monitoring procedures surrounding their use (Adams and Jennison 2007; Lin and

Jones 2010). There is also no definition of exactly what *use* of a LLW implies. It is not agreed whether the display of a LLW and/or a threat to use or fire such a weapon constitutes *use* of that weapon.

2.13.3 Reactions and responses to less-lethal weapon problems

The primary response to LLWs has been the controlling of their use via national regulation; the result being a particular LLW can be banned in one country but permitted in another. This is a political decision each individual state needs to take.

If a particular LLW receives political approval then possession and use of that LLW depends on the domestic laws of the individual countries. In some jurisdictions civilian possession and use of pepper spray and/or Tasers is allowed; for example under Sections 12650-12655 of the California Penal Code, any person may purchase, possess, or use a stun gun (i.e. Taser) providing they do not fall within a number of restricted categories (i.e. drug addicts, under 16's, those with felony or assault convictions, etc.). Whereas in other countries civilian possession of these items is restricted; for example within the UK the purchase, possession, manufacture, use, or sale of a Taser is prohibited under s.5(1)(b) Firearms Act 1968. However, UK police are permitted to carry Tasers on completion of a three-day training course (London Assembly 2013; Aiming for a safer solution 2013).

In other instances a state may only sanction the use of certain LLWs in specific instances, such as for mass crowd control. These can include measures such as water-cannons, CS gas, and baton rounds.

2.13.4 Identified controversies arising from less-lethal weapons

Table 2.12: Identified controversies arising from less lethal weapons

Code	Technology = **Less Lethal Weapons**
LW1 *LW1a*	Two conflicting purposes/justifications for their use, with support often dependent upon which purpose/justification is presented as the rationale for introducing these weapons; • (1) raising the threshold for the use of deadly force by providing an

LW1b	alternative means to subdue a person, and • (2) Providing tools of compliance to make citizens obey police orders out of fear of pain and/or incapacitating them.
LW2	Once introduced, there is *function creep* pressure to employ LLWs for purposes other than as an alternative to lethal force.
LW3	Tasers are too readily used by police to force people to comply.
LW4	There are ethical and societal questions about police using pain causing devices like Tasers to force people to comply.
LW5 *LW5a*	The use of LLWs in inappropriately situations and for inappropriate purposes can undermine the continued use of such weapons; • Using crowd control LLWs against political protestors undermines democracy and individual rights (e.g., the use of cater cannons against civil rights protestors in 1960's USA potentially means they will never by acceptable in the US again).
LW6	Unrestricted use of LLWs against citizens will convert them into enemies.
LW7 *LW7a* *LW7b*	LLWs can and do kill, even when operators follow proper operating procedures and guidelines; • Manufacturers and agencies who deny this fact lose credibility. • Manufacturers and agencies who deny this undermine the credibility of the LLW.
LW8	LLWs can and do cause permanent injuries even when used correctly.
LW9 *LW9a*	Lack of thorough pre-deployment evaluation and safety studies by independent researchers denies a LLW a measure of legitimacy. • Positive research pertaining to a LLW paid for by the manufacturer is open to questions of bias.
LW10	Variability in output between units of the same LLW, whereby they produce stronger effects than a unit acting according to manufacture specifications undermines the perception of that LLW and may increase the risk to safety of that weapon.
LW11	The question manufactures should be addressing by their products is not *are they safe?* Rather *is their weapon as safe as, or safer than, the alternatives?*
LW12	Individual differences in targets which increase the risk of injury and death from LLW use, and which may not be apparent to the user of the LLW represent the single biggest challenge for LLW designers; This includes amongst others;

LW12a *LW12b* *LW12c* *LW12d* *LW12e*	• Those on medication or illegal drugs. • Asthma sufferers. • Heart conditions. • Mental conditions. • Pregnancy.
LW13	Certain categories of people may not be appropriate targets for LLWs, such as pregnant women, children, elderly people, those on drugs, and those with medical conditions.
LW14	Testing of LLW prototypes can injure the subjects and pose a risk of death.
LW15 *LW15a* *LW15b*	Area effect LLWs such as malodorants and sticky materials may require decontamination of targets and the physical environment, and may also hinder police; • They may also make it impossible for police to effect arrests without protective clothing/equipment. • Obscurants like smoke grenades affect police trying to make arrests.
LW16	Police are using Tasers as weapons of punishment.
LW17	Tasers have been used on psychiatric patients in hospitals which is condemned by authorities and medical practitioners.
LW18	Electro-shock weapons are used as torture devices where the absence of lasting marks makes them particularly useful.
LW19	Use of LLWs on prisoners in jails risks overuse with little accountability.
LW20	Shooting unarmed suspect with Tasers for arguing or talking back, walking away, or being discourteous undermines public support for these weapons and the people who use them.
LW21	LLWs have been used on ethnic minorities in situations which give rise to claims of police racism.
LW22 *LW22a*	Overuse of LLWs subverts traditional policing practices by discouraging non-violent methods of persuasion. • Using LLWs in pre-emptive manners to induce compliance might lead to charges police being *trigger-happy*.
LW23 *LW23a*	LLWs can increase the *use of force* instances by police when used to produce compliance in situations where physical force would not be justified. • Using rubber bullets with the inherent risk of serious injury/death in situations where real bullets would never be acceptable undermines

LW23b	the credibility of use. • LLWs cannot substitute good training, judgement, policies, procedures, and leadership within the police services.
LW24	Citizens engaged in protests are demonstrating a grievance. Use of force via LLWs which injures these protestors may disperse them but the grievance remains.
LW25 *LW25a*	A LLW cannot survive without public support; • Unacceptable facts or publicity will affect the public and politicians.
LW26 *LW26a*	There is no definition of what *use* of a LLW denotes for compiling statistics. If they are used as tools of compliance then threatening to use a LLW, removing it from a holster, or pointing it at a person may all constitute incidents of *use*. • Failure to correctly record the use of LLWs may hide abuse.
LW27	Arguments have been made that LLW use should require both a *subject matter threshold* (such as actual bodily harm) and an *immanency requirement*. Thus acts of common law assault or walking away will not be sufficient grounds for deploying LLWs.

2.14 Coding of Results

The controversies identified in Chapters 2.2-2.13 are now collated and individually categorised by both the *origin* and the *nature* of the controversy[103]. The results of this process are presented in Appendix *E*. By *origin of controversy* I am referring to the generalised point in the process of creating a ST at which any specific controversy is seen to originate; starting from the process of physically designing that which was previously an idea or set of specifications, through to implementation, end-user use, and the on-going oversight of that technology. *Origin* is divided into three groups;

1. *Design Features*: controversies originating from and/or in response to aspects of the physical/digital construction of the security technology. This includes not just what the technology is *designed to do* but also aspects such as the outputs

[103] The coding categorisations I have employ within Chapter 2.14 are the result of an initial analysis of the raw data collected through the case studies undertaken in Chapters 2.2-2.13. However, this remains a subjective process and another researcher analysing the same data may well settle on alternative categories.

produced by this technology (such as CCTV footage, the electrical discharge of a Taser, the digital image from a body-scanner, and the output of an algorithm), the physical appearance of the technology, etc.

2. *End Users*: controversies arising from the actions of end-users (those actively employing the controversial security technologies). This group includes both state employees (police, local council workers, etc.) and private citizens (shopkeepers, corporate employees, etc.); restricted only by the existence of any technology-specific rules governing the employment of a particular security technology.

3. *Policy Decisions*: a broad category incorporating controversies which derive from *official* decisions as to how a technology will achieve its intended security goals, usually made by public officials (politicians, senior police, etc.) or those operating highly regulated industries with national security implications such as airport operators. This category feeds into both *Design Features* and *End Users* but exists either above or a step removed from these other two categories. Using an automated threat detection algorithm as an illustration; while the *black-box* that is the algorithm itself is a *Design Feature*, and the actions of those monitoring or acting in response to the algorithm-output represents *End Users*, the decisions to create/employ the algorithm and whether or not the final variables will be made public, rules as to what variables can/cannot form part of the algorithm, and the creation of codes-of-practice for end users are all *Policy Decisions*. Given the breadth of this category I have divided it into two sub-categories;

    i. *Rules Governing Use*: focuses specifically on official/unofficial guidance governing how end users shall use the technology and/or respond to any output.

    ii. *Wider Policy Decisions*: all remaining *Policy Decisions* not covered by the *Rules Governing Use* sub-category.

By *nature of controversy* I am focussing on the specific fundamental element(s) which underlie and define each of the individual controversies identified in Chapters 2.2-2.13 above. Eight *nature of controversy* categories were identified from a preliminary examination of the twelve controversial ST case-studies. These eight categories are

scrutinised further to identify sub-categories, as represented by the '*Key words, phrases, concepts*' column of Appendix *E*. The eight identified common categories representing the *nature* of the controversies examined are:

1. *Health*: controversies arising from real, perceived, or unknown health risks resulting from the security technologies in question.
2. *Legality*: controversies arising from the potential or determined illegality of the security technologies as determined under national and international law. Note however that this section does not include potential or determined breaches of the European Convention on Human Rights as incorporated into UK law under the Human Rights Act 1998 which are covered under the *Rights & Liberties* category.
3. *The Public*: controversies based on the public's views, opinions, and support for a security technology. Also controversies arising from the effect of the security technology on both the whole of society as well as minority groups within it.
4. *Rights & Liberties*: controversies based on the perceived detrimental effects of the security technologies on the concomitant rights and liberties of individuals, groups, and the whole of society. This category includes potential or determined breaches of the European Convention on Human Rights as incorporated into UK law under the Human Rights Act 1998.
5. *Cost*: the financial cost of a security technology resulting in controversy.
6. *Safety & Security*: when a security technology directly or indirectly jeopardises, potentially jeopardises, or is perceived to jeopardise, the safety and security of citizens in the course of its operation.
7. *Functionality*: controversies based on features and aspects of the actual design of the security technologies in question (such as whether or not the security technology works, its reliability, the proportionality and intrusiveness of its feature, etc.).
8. *Use & Misuse*: controversies arising from the misuse of, or propensity for misuse/abuse of, security technologies.

2.14.1 Discussion of coding results

There are a number of observations which can be made from the results of the coding process presented in Appendix *E*. These are as follows:

1. *The same controversy can arise repeatedly in relation to very different STs*: forty-three common controversies were identified from the case-studies, the vast majority of which appearing in more than one of the twelve STs examined.
2. *Categorisation of commonalities is possible*: aided by the recurring nature of the controversies identified, it is possible to objectively organise those identified into categories and sub-categories (see Appendix *E*).
3. *Despite the objective elements of this categorisation process, it remains underpinned by subjectivity*: the 7 categories employed within Appendix *E* to arrange the 43 identified commonalities does not represent the only possible configuration for arranging these commonalities. For example categories 2 *Liberties and Human Rights* and 3 *Questions of Legality* could easily have been combined, and/or separate categories for *data protection* or the issue of *trust* could also have been included. Also, at the commonality level cases could be made for either reducing or increasing the number of commonalities from 43 by combining or breaking-apart those included. As these decisions remain at the discretion of those involved in producing any taxonomy, subjectivity remains. However, *subjectivity* does not mean *totally arbitrary*. I content that the categories chosen successfully incorporate all 43 identified commonalities while at the same time being sufficiently distinct, coherent, and largely recognisable to a first-time reader, that they constitute logical divisions.
4. *I had not anticipated all of the 43 commonalities identified before beginning the case-study process*: while certain commonalities (such as those focussed on privacy infringement, direct discrimination, or illegality) were not unexpected based on my previous research projects, others were far from obvious to me. Commonalities such as: *3c. Citizens cannot determine if the technology is operated legally*; *4c. The financial figures released into the public domain are not trusted*; and *7c. The presence of the ST potentially jeopardises the safety and security of citizens*, were ones I had not anticipated. This is an important point as it gives

weight to the argument that potential triggers for controversy are not always self-evident during the design phase and before they manifest in the form of a social resistance. Thus those developing new STs may require assistance in identifying these triggers.

5. *The commonalities identified are not absolutes, rather they are based on spectra and balancing*: few of the 43 commonalities represent binary acts or events whose simple presence can result in controversy[104] with one possible exception being *1f The ST causes a fatality*. Many of the commonalities are based-on/incorporate a spectrum; for example *4b. The cost of the technology is considered too high or excessive by the public* (the implication being that the probability of avoiding controversy will be increased if the ST can be delivered and operated for a lesser price), and *5d. The ST is not necessary* (the implication being that should circumstances change such and public opinion hold that the ST has become necessary then the probability of avoiding controversy will increase). Others imply the search for a balance between what is acceptable and what is not. Consider *5f. The ST represents a disproportionate response* and *6e. The ST is ineffective and/or incapable of addressing the identified security problem(s)*; these two identified commonalities can be interpreted as two ends of a spectrum (disproportionate response versus ineffective response) implying that arriving at an acceptable ST will require developers/designers achieve an acceptable balance.
6. *The importance of 'trust'*: a recurring theme which crosses the seven constructed categories within Appendix *E* is that of *trust*. It is explicit in a number of the 43 commonalities[105], and for many of those remaining it is arguable that without public trust the likelihood of a commonality evoking social controversy increases.

[104] At no point here (or anywhere else in my dissertation) am I claiming that the presence of a commonality of controversy *will* result in social resistance to a proposed ST. Rather within point *5* I am discussing the different natures of the identified controversies (i.e. in what form they manifest). Even if they do appear within a ST, regardless of how the particular controversy arises, there is no guarantee controversy will follow as a result.

[105] See 1g, 3c, 4c, 5b, 5g, 6b, 7a, 7b, 7c, and 7e within Table 2.14.

2.15 Taxonomy of Security Technology Controversies

By combining the keywords/concepts identified in Appendix *E* it is possible to identify 43 recurring concepts; referred to throughout this document as *commonalities of controversy*. To provide order these have been grouped into seven categories[106], as described below. Each is accompanied by both a short statement describing the central focus of that category, as well as a number of bullet-points defining its breadth. This is followed by a depiction of the completed taxonomy on page 155 which incorporates the seven categories as well as the 43 commonalities of controversy.

1. PHYSICAL OR MENTAL HARM

The ST possesses the potential to cause physical or mental harm to an individual.

- This harm does not need to be serious or life-threatening; for example any form of pain, discomfort, nausea, dizziness, physical or mental impairment will suffice.
- This harm does not need to be likely; statistically rare events of harm will suffice.
- The effects do not need to be permanent. It includes transient harms such as short episodes of nausea, pain or dizziness.
- It is irrelevant whether this harm will only arise through either the misuse of the ST or some form of accident/failure.
- While the harm may be rare it must be plausible; i.e. using a body scanner as a weapon by crushing someone with it is not a plausible harm, however using a Taser-style baton as a club in close-quarters is a plausible harm.

2. LIBERTIES AND HUMAN RIGHTS

The ST possesses the potential to infringe upon other human rights when seeking to provide security.

- Human rights of likely concern here include (but are not limited to):
    - privacy;
    - informational self-determination;
    - assembly;

[106] These are based on the eight *nature of controversy* categories applied within Table 2.13, with two of these categories ('safety & security' and 'use & misuse') combined into one (i.e. Safety, Security, Misuse & Abuse).

    - association;
    - freedom of expression;
    - right to a fair trial;
    - freedom from torture or inhumane or degrading treatment.
- The provision of security does not automatically outweigh or *trump* the enjoyment of competing rights.

3. QUESTIONS OF LEGALITY

The potential exists for the legality of this ST to be questioned, challenged, or brought into doubt.

- This includes situations where: STs raise new issues of law where precedents have not been set; where the technology has not faced legal scrutiny; and where it is possible for somebody to evoke existing legislation to challenge its legality.
- STs which possess the power to discriminate, or be used in a discriminatory manner are included herein.
- STs which engage data protection principles are included here.
- If the potential exists for the ST to be used in an illegal manner, or if this technology has been deemed illegal, subjected to bans or restrictions of use, then they are included in this category.

4. FINANCIAL COST OF THE SECURITY TECHNOLOGY

Given the financial cost of the ST, it does not represent a sound investment.

- 'Financial costs' go beyond the purchase price; they also include any on-going operation, training, maintenance, calibration and upgrade costs, as well as any one-off charges (such as changes to building layouts, etc.).
- This questions whether the security benefits of a ST outweigh the financial outlay.
- If similar security gains can be achieved through other less expensive means, then a ST is called into question.
- Public disclosure of costs is a factor within this category.

5. PUBLIC AND END-USER ACCEPTABILITY

Issues undermining public and/or end-users support for a ST.

- Covers the issue of gaining and maintaining *trust* within a ST from the public and end-users. Also questions whether support for a ST is conditional.
- Asks whether a ST can be justified, whether or not it is proportional and whether it meets social acceptability standards?
- Views the technology through the eyes of a society to determine whether it meets recognisable values possessed by that society.

6. ISSUES OF FUNCTIONALITY

The functionality of a ST allows it to be criticised or called into question ('functionality' of a ST encompasses; what it does, what it doesn't do, what it's intended to do, and what it is capable of doing).

- Functionality represents the culmination of all the individual functions (i.e. the individual components/ abilities/modes-of-use/ capabilities/etc.) built into the ST. These should all be necessary, proportional, and reliable.
- The maturity/reliability of the science/technology underlying the ST is examined here.
- Both the operation and effectiveness of the ST are examined.
- If the ST suffers from errors, or is susceptible to function-creep or dual-use, this will lead to questions over its functionality.

7. SAFETY, SECURITY, MISUSE AND ABUSE

The ability of a ST to be misused or abused thereby jeopardising the safety or security of citizens or their property, as well as the ability for attackers to avoid or circumvent that ST.

- Includes past misuse/abuse of similar STs.
- If a ST is only safe and/or secure if used in accordance with the manufacturer's instructions then the potential for misuse or abuse cannot be ruled out.
- Abuse or misuse of a ST can be by governments, police, state officials, insiders, end-users, businesses, external attackers, and citizens.
- The ease with which a ST can be avoided or circumvented is examinable here. As is the propensity for a ST to become a *honey-pot* for attackers.

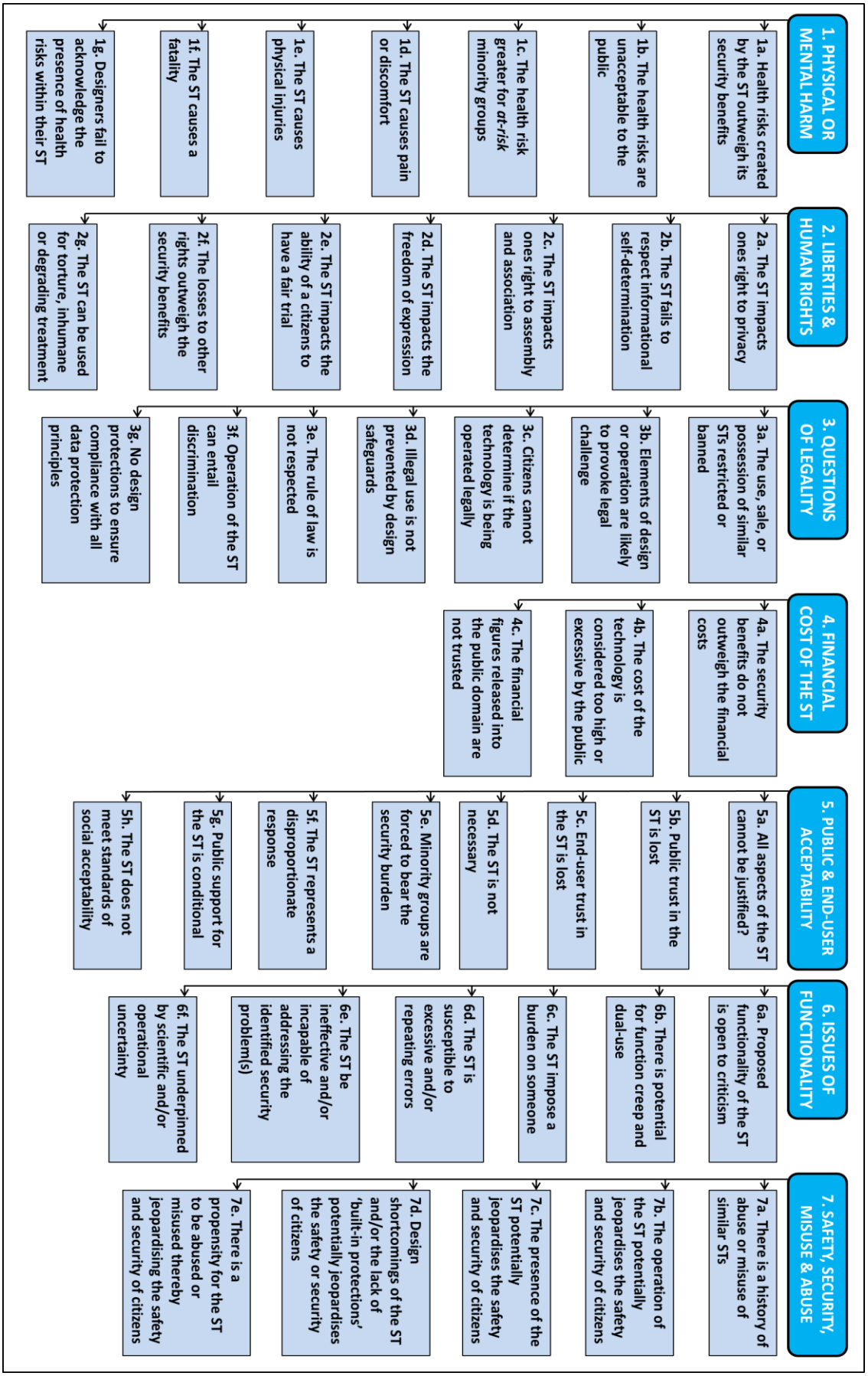


Table 2.13 Completed Taxonomy of Security Technology Controversies

3. Stage 2 – Initial Interviews with Engineers and Scientists

Following the workflow illustrated in Figure 1.1 Macro-level structure for conducting the research project, this Chapter constitutes Stage 2 *Interviews with Engineers and Scientists*. The goal here is to gain an understanding of how STs are produced by interviewing those engineers and scientists with experience in actually developing these technologies. Areas focussed on included:

- the education of these developers,
- how a ST moves from being an idea or tender through to becoming a completed technology,
- the roles of individual developers/designers within this process,
- and how any future design tool for mitigating negative social reactions to ST would need to look and operate.

The overarching goal of Stage 2 is to provide information that can be combining with the output from Stage 1. Combined, these will form *requirements* for any future design tool; a process to be undertaken in Chapter 4.

3.1 Methodology

As outlined in Chapter 1.1.1 above, the purposes of Stage 2 is to develop an understanding of how the designing and developing of STs occurs in practice. In turn this requires gathering knowledge on, and from, the individual designers involved in this process, including their motivations, their skill-sets, and the constraints inherent to working within the ST industry. To achieve these goals the methodology I am employing is *semi-structured interviews*.

Semi-structured interviews are described by Longhurst (2010, p.103) as:

> a verbal exchange where one person, the interviewer, attempts to elicit information from another person by asking questions. Although the interviewer prepares a list of predetermined questions, semi-structured interviews unfold in a conversational manner offering participants the chance to explore issues they feel are important.

This inbuilt degree-of-flexibility allows the researcher to tailor each interview to the interviewee so as to maximise the value of their responses, by "*adjusting the level of language of planned questions or through unscheduled probes … that arise from the interview process itself*" (Berg and Lune 2012, p.112). For my research project this flexibility is particularly useful for meeting the goals of Stage 2 for two reasons. Firstly I have already identified the general areas of information I wish to collect (i.e. ST development processes and the interviewee's skill-sets) but lack first-hand knowledge of this area, so semi-structuring permits the exploration of responses while keeping the overall interview sufficiently focussed. Secondly, while the interviewees share many characteristics, in other aspects they differ greatly. They are on the one hand homogenous in that they are all STEM practitioners working on the development of STs, while on the other hand they are heterogeneous, possessing different levels of seniority, responsibilities, experience, education, and they work in different environments.

There are numerous strengths to interviewing which mesh well with both the data I need to collect and my interviewees with their particular need for anonymity. These strengths (and the associated benefits for my research project) include the following:

1. Interviews can fill knowledge-gaps more efficaciously than alternative methods such as observations or the analysis of documentary data (Hay 2010). Having already identified commonalities of controversy in relation to the finished design of STs through multiple case-study analysis in Stage 1, interviews can fill in gaps surrounding how individual STEM practitioners actually go about designing and building these STs.
2. They enable the investigation of complex behaviours and motivations (Hay 2010) and the elicitation of 'deep' answers to questions (Guest et al 2013). Given the complexity and diverse nature of the ST industry, interviews provide a perfect vehicle for data collection here.
3. They can reveal both consensus on issues and diversity of opinions and experiences (Hay 2010). As touched on above and examined in depth in Chapter 3.2.4, both the diversity of the individual interviewees' experiences and the

uniting factor that is their occupation as ST designers, can be effectively examined through the use of interviewing.

4. Interviews respect and empower the interviewees, while also providing them with more information on the research project they are participating in than some other methodologies (Hay 2010). In the context of my research, interviews allow me to respect the anonymity wishes of interviewees who are naturally cautious about participating in research. Additionally, because interviews are an interactional process whereby information flows between the interviewer and interviewee thereby enabling the interviewee to develop a deeper understanding of the research project, I am better able to obtain specific, reflective feedback regarding future design tools.

On the basis of these factors, semi-structured interviews constitute an excellent methodology for successfully completing Stage 2. Alternate methodologies that could be employed here include in place of interviewing include questionnaires, surveys, focus groups, and ethnographic research. However, these are rejected as being less effective, unsuitable, or impractical.

The effectiveness of questionnaires and surveys is undermined by the small number of willing subjects and the difficulties involved in accessing more. Also, the detailed nature of the information I need to obtain from my participants, which itself changes depending on the particular characteristics of each participant, negates the value of producing and employing standardised collection tools such as questionnaires/surveys. Focus groups, while potentially just as effective as interviewing for meeting the goals of Stage 2, are necessarily excluded due to the anonymity pre-conditions insisted upon by a number of the participants. Finally, ethnographic methods are not considered a viable option for this research project as I do not have the necessary long-term access to those workspaces where STs are being designed and developed.

3.2 Method

3.2.1 Constraints on conducting the interviews and publishing the results

Before I expand upon the interview process and present the results collected therein, it is prudent to explain the prerequisite constraints agreed with the interviewees prior to their interviews and which have impacted how the results of this process are presented within this chapter. Ultimately these constraints constitute safeguards for protecting the anonymity of the interviewees.

As discussed in Chapters 1.3 & 1.4 the research and design of STs conducted within the security industry often carries with it a number of unique challenges and restraints. These restraints are primarily a result of the secrecy requirements affecting much of the work in this domain. This secrecy can be propagated and enforced using commercial-secrecy justifications to prevent discussions of patents held and work undertaken, employment contracts which can include gagging provisions, and statutory legislation. Regarding legislation, within the UK the Official Secrets Act 1989 applies to both Crown servants[107] and government contractors[108] such that the disclosure of that information protected by this Act (such as the design of STs) may constitute an offence[109]. This includes information which if disclosed will *impede the prevention or detection of offences or the apprehension or prosecution of suspected offenders* (as per s.4(2)(a)(iii)). To provide context for how this offence may arise in relation to the design of STs, consider the example of whole body scanners examined in Chapter 2.2. The successful detection rates of these scanners in relation to plastic and liquid explosives hidden under clothing (as determined by both government testing and any detection figures collated from in-field use in airports) represents protected information which has never been publically disclosed. While this secrecy impedes public debate, makes it harder to gain public trust, and provides the basis for controversy, using the language of s.4(2)(a)(iii) Official Secrets Act 1989 the decision to

[107] 'Crown servants' are defined under s.12(1) Official Secrets Act 1989 as including Ministers of the Crown, civil servants, police, military personal, and others.

[108] 'Government contractors' are defined under s.12(2) Official Secrets Act 1989 as including non-Crown servants who are employed to provide goods and services to those who fall within the category of Crown servants as set out within s.12(1).

[109] The maximum penalty here being 2 years imprisonment for conviction on indictment as per s.10(1)(a).

keep this information secret is justified by decision-makers on the basis that these figures possess the potential to impede:

a) *the prevention of offences*: if the actual detection rates are low then the scanners will lose their deterrence effects;
b) *the detection of offences*: if detection rates differ depending on the substances carried, the amount of this substance, where it is carried, etc., then attackers can modify their carrying methods to maximise their probability of becoming a false negative, thereby reducing the scanners' detection capabilities;
c) *the apprehension of suspected offenders*: a reduction in the detection of offences will have the concomitant effect of reducing the apprehension of offenders.

It is clear from this discussion that those involved in the production of these technologies are necessarily apprehensive about discussing their work. To garner the cooperation of the interview subjects I needed to agree to all their preconditions to involvement; these differed markedly between interviewees. Some were either happy or ambivalent as to being identified while others demanded anonymity, so to proceed I collated the various requirements from all the participants and applied the most onerous versions of each of these to all the data collected. This approach allows me to retain the maximum number of subjects for interviewing. The following list of restrictions is the result of this process:

- Anonymity for the interviewees.
- The inclusion of measures to avoid or impede the identification of interviewees through the aggregation of their quoted responses. This was achieved via two means:
    1. By not assigning any identifier to any of the quotes included (such as *subject A*, *subject B*, etc.).
    2. The removal or modification of references to specific commercially-identifiable STs which the interviewees personally worked on. As a fictitious example; if *Subject A* was employed by Taser International in the development of the TASER X26P[110], I would remove any reference they

[110] See Chapter 2.13 for the case-study on less-lethal weapons.

made to their specific employer, and would substitute a generic moniker for the product – i.e. "*a stun-gun*".

This concern over the potential to identify individuals through data aggregation and the lack of trust in pseudo-anonymisation is justifiable based on research from various fields (see amongst others: Malin and Sweeney 2004; Sun et al 2012; Aimeur et al 2012)

- The permanent deletion of all recordings and subsequent transcripts of the interviews after analysis to prevent non-consensual secondary analysis and function creep.
- The opportunity was provided for interviewees to vet this section before publication.

These restrictions do have a negative effect on the qualitative process of reporting interviewee quotes. Mainly in that it prohibits both: i) the joining together of quotes from a single interviewee to form a more detailed picture of that individual; and ii) comparing the views of one individual with others in the study. However, I believe the value and contemporary relevance added to this research project by being able to interview engineers and scientists who are directly engaged in the process of designing and developing STs far outweighs that which is lost.

3.2.2 The interview process

The three primary research questions/areas being examined within the interview process of Stage 2[111] were the following:

1) Developing an understanding of the professional education of the interviewee, with specific emphasis on determining what (if any) social science and engineering ethics training was both offered and undertaken.
2) Examining how a ST goes from being an idea and/or prospective tender to a research and design project and finally to a finished product; including the divisions of labour, responsibility, knowledge, and power within a design project.

[111] See Figure 1.1 Macro-level structure for conducting the research project.

Also to determine what 'social acceptability' factors influence this process (if any) and how this occurs.

3) Collecting what are essentially end-user views on what they see as the optimal design, format, and operation of practical tools for influencing the upstream development of STs so as to maximise their social acceptability.

Obtaining *first-hand* information from the designers and developers of STs on these three areas was the primary motivation for engaging in this interview process.

The interviews undertaken during this Stage were conducted in the form of *semi-structured interviews/guided conversations*, as distinct from the more polar alternatives of either; (a) highly structured, standardised interviews, such as survey interviews with their quantitative orientation, or (b) unstructured interviews akin to free flowing informational exchanges where the interviewee directs the interview (Holstein and Gubrium 2004; Longhurst 2010; Dunn 2010).

Semi-structured interviews were chosen so as I could focus the conversation on specific topics (Rubin and Rubin 2005) pertinent to the three primary research questions/areas listed above. While at the same time emphasising the content of what the interviewees have to say (Dunn 2010) by affording them the opportunity to express themselves beyond simple 'yes'/'no' answers (Longhurst 2010).

To facilitate this process an interview schedule was produced containing fully worded questions often accompanied by possible follow-up's[112]; however the interviewer was not bound to ask *only* these questions, nor *all of* these questions. Based on factors including; the answers provided by the interviewee, their knowledge, employment position, experience, etc., the interviewer could modify the questioning to explore different areas of interest as they arose. Additionally it was the interviewer's responsibility to redirect the conversation back to the pre-identified research-topics/questions when this was the appropriate action.

[112] See Table Appendix *F*

3.2.3 The Stage 2 questionnaire

Appendix *F* depicts the selection of questions posed to the interviewees. This question schedule is divided into four sections:

- *Education questions*: focussing on the interviewee's professional qualifications with dedicated questions to determine their social-science and ethical education.
- *Working practice questions*: focussing on how they actively produce STs within their workspaces, how they interact with co-workers on projects, how work is divided and assigned, and their ability to influence designs.
- *Thoughts and opinions for design tool*: these questions are designed to illicit opinions on what shape future STs should take (i.e. how they should operate, their output, how long they should take to use, etc.).
- *Miscellaneous*: asks for any information not covered above which the interviewee deems relevant based on the exchange of information inherent to the semi-structured interview process.

The question schedule is divided into lead questions which are assigned an identifying letter/number combination (e.g. A1, B2, etc.). Each of these is linked to potential follow-up questions where appropriate.

3.2.4 The interviewees

To enhance the credibility of any interview-based research, Rubin and Rubin (2005) recommend selecting interviewees who:

a) Possess first-hand knowledge of the research problems being examined.
b) Ideally are speaking from first-hand experiences.
c) Are individuals who reflect a variety of different perspectives.

Within the research project, fifteen individuals were interviewed during this stage. When compared against the criteria of Rubin and Rubin; all of these individuals are either currently or previously employed in the research and development of STs. This affords them first-hand knowledge and experience in the research topics being examined. They also reflect a variety of different perspectives, though these can be

conceptualised as perspectives *internal* to the ST industry as opposed to *external* ones. In other words, even though all fifteen interviewees work(ed) within the ST industry they possess different perspectives given that; (i) the ST industry is vast and varied, and (ii) all fifteen subjects possess different combinations of experience, current/past employment positions, and university degrees/specialisations.

I chose not to include similar individuals (scientists, engineers, etc.) from outside the ST industry (whose perspective will necessarily be *external* to the ST industry) because;

a) this research project is not a comparative study of the work practices between different science and technology industries, and
b) as discussed in Chapters 1.3 & 1.4 the unique nature of the ST industry sets it apart from many other science and technology endeavours. Hence the decision to focus only on those with first-hand knowledge of this industry.

Regarding the diversity of the sample of interviewees, and without providing specific numbers to assist aggregation of personal data, the fifteen ST designers collectively possessed the following characteristics:

Table 3.1 Diversity of interviewees within Stage 2

Academic Qualifications: • Electronic engineering • Structural engineering • Software engineering • Medical physics • Physicist • Mathematician • Explosives engineering
Employers: • University • Self-employed /company director • Government • Multinational ST company • Domestic ST company • Contractor
Current employment role: These employment roles covered the full spectrum from:

- Junior research associates/engineers through to chief technical officers
- The junior employees of an organisation through to the owners/directors

The selection process for this Stage involved *non-probability sampling*; in that only those individuals with experience in developing and designing STs were considered eligible for inclusion. This form of sampling is described by Lewin (2005 pp.218-19) as:

> … [an approach] adopted when researchers target a particular group and are not always seeking to generalize findings to the population overall. This kind of approach is commonplace in small-scale research (particularly when costs need to be minimized) or qualitative approaches….

Non-probability sampling is widespread within social science research (Lewin 2005). It was particularly appropriate here given my focus on; (i) only targeting those scientists/engineers engaged in the process of designing STs, and (ii) my recognition that any findings within this research project are not intended as generalisable beyond the ST industry due to its unique nature.

The particular forms of non-probability sampling undertaken were a combination of:

a) *Convenience sampling*: whereby participants are selected based on ease of access[113] (Bradshaw and Stratford 2010; Lewin 2005).
b) *Snowball sampling*: "*A technique used by researchers whereby one contact, or participant, is used to help to recruit another, who in turn puts the researcher in touch with another. The number of participants soon increases rapidly or 'snowballs'*" (Clifford et al 2010 p.535).

Obtaining initial access to prospective subjects was difficult given the nature of the ST industry (see Chapters 1.3 & 1.4) and the concerns-of/restrictions-on employees within this industry (see Chapters 1.4 & 3.2.1). To overcome this I began with convenience sampling by exploiting both my personal networks with practitioners, and those networks existing within UCL's Department of Security and Crime Science to identify the initial cadre of willing participants. From this group I then employed snowballing by asking interviewees to introduce me to other potential subjects. Through this I was

[113] Though it should be noted here that *ease of access* is a relative term – the use of personal and departmental networks represented the easiest, most convenient method for gaining access to subjects.

able to conduct fifteen interviews with engineers and scientists possessing direct experience working in the design and development of STs.

The interviews varied in length from approximately thirty minutes to four hours; however, these figures do represent outliers. The majority of interviews lasted around sixty minutes. Face–to-face interviews were undertaken throughout the United Kingdom and Europe. Interviewees outside these areas were interviewed via Skype. All interviews were recorded with a digital recorder before I transcribed them into scripts. Coding of these scripts was then undertaken manually so as to identify the presence or absence of common themes in relation to each question.

While it would have been possible to conduct more than fifteen interviews, two factors truncated this process. Firstly, given that this interview component comprised only one stage of a larger research project, unlimited time could not be afforded to it. Secondly, after analysing these fifteen interviews it was determined that an identifiable *saturation point*[114] in responses was being observed justifying the decision to forego further interviews.

3.3 Results

The following are selected results from the Stage 2 questionnaire[115]. The presentation of these results is governed by the restrictions described in detail in Chapter 3.2.1. Where there was an identifiable common theme or view in relation to a specific question, this will presented here, and they may be accompanied by unassigned quotes when this adds value. Given the semi-structured nature of the interviewing, not all of the questions set out in Appendix *F* were posed to all of interviewees (see Chapter 3.2.2 for more details), and as a consequence not every question provided suitable results for inclusion in the section below. Additionally as the interviewing process was designed to enable the examination of interesting concepts as they arose (including those beyond the scope of the original questionnaire), additional questions and their results are also included here.

[114] Denoting the point of diminishing returns whereby new responses or concepts are no longer emerging within new interviews (Corbin and Holt 2005; Rubin and Rubin 2005).
[115] See Appendix *F*

3.3.1 Responses to Stage 2 questionnaire questions

Education questions:

- *With regard to the composition of your university degree, was it rigid and structured or was there flexibility in the choice of subjects you could take?*

In responding to this question all the interviewees characterise their courses as rigid in structure regardless of their chosen discipline. However, there are two other themes which expand upon this characterisation. The first is that this rigidity is front-loaded in their chosen degrees in that there was little or no option to take elective courses in their first year. However, as their degrees progress the opportunity to select elective subjects increases (response: "*we had a choice of 1 module in the second year and 3 modules in the third year .. and they were all engineering courses, as in hard-core engineering. I didn't consider them a big choice*"). Though, in deciding whether or not a person has *choice* encapsulates a subjective determination. As the final sentence in this response indicates, the opportunity to select subjects is not necessarily equated to 'freedom of choice' by the student if they perceive the options to choose from as largely homogenous.

Secondly, as their degree progresses the opportunity opens up to specialise in areas which most interest them, though these specialised subjects are still built upon a core of prerequisite subjects. This opportunity to choose subjects is characterised as entailing both flexibility *and* restriction: flexibility in that it allows for greater choice, but restrictive in that once a specialisation path is chosen it becomes harder to change direction.

- *Were there any compulsory or elective subjects that you would describe as social or social-science subjects, such as the role of the engineer in society, the responsibilities of engineers, and/or the impact of engineering on society?*

The vast majority of respondents emphatically answer 'no' here. However, two engineers do indicate that their degrees include a level of ethical training. One holds this to be implicit to their education (response: "*all engineers learn ethics. When we're*

taught to build a bridge you learn how to build it so it doesn't fall down in five years from now. That's being ethical"). The other identifies an explicit subject undertaken which teaches ethics (response: "*on my MSc we had a course called xxxxx that dealt with the contractual and the public interaction of engineers [when] bidding for a job and that had ethics, engineering ethics, into it because that's where we would put it*").

- *If there was something offered, would it have interested you?*

Responses here are generally non-committal, with 'maybe' being was the most common offered. However, there is not a rejection of social science or ethics displayed by the interviewees. In discussing this question, they will often caveat their response by adding that they would consider taking any social-science/ethical training providing they consider it to be useful. The following statement is a good example of such thinking (response: "*I don't think so no because the course was very much about maths*" … {and then later} … "*had they said 'okay were going to look at the social side of security' then maybe I would have been interested in that, had I known that I wanted to go into [the security] area*"). That said, this example also displays the difficulty in applying this logic of *useful social science training being acceptable*. For a student will not know exactly what *work* their future career will entail, so is not going to know what training is going to be useful to them in their future when they are selecting their subjects.

<u>Working Practice questions:</u>

- *What is the process whereby you decide whether or not to take on a new ST project put out to tender by the government?*

Those interviewees who were in more junior roles, or had never been exposed to this process because of their employment role, were unable to answer this question, citing a lack of knowledge of this process. Those in more senior positions working in non-government roles were able to address this question. In their answers the single factor which determined whether or not they would even consider tendering for a project was the perceived ability to make a profit from any final product. This factor was the

primary, determining one. It was addressed and used to filter out possible projects even before considerations over the technical ability of the firm to deliver the tender were considered. The following response details how this process is undertaken by one of the interviewees with considerable experience in this area.

> (response: "The first thing that happens is it [the tender] goes to marketing. They talk to our clients, to potential end-users. And they come back and tell us, you know, if we actually build this thing, whether there exists a market for it. Can we sell enough of it to make a profit. If they [marketing] say 'no' then we don't touch it [the tender] …. If they say 'yes' then the Chief Technical Officer (CTO) picks it up and [he/she] have to decide whether or not it's feasible …. Do we possess the capacity, the capabilities, to build this thing? Also is this the sort of thing we [the company] want to be building. After this [he/she] will bring in the heads of the different departments to look at how it might be built")

Despite working in the national security area, this description of the tendering process indicates that private companies make their decision based on monetary factors rather than a perceived nationalistic duty.

- *Explain what happens when a new design project is announced?*
  - *How is the work for these projects divided?*

There was no uniformity of process displayed here in relation to how the work within projects was divided up, beyond the fact that *somebody* sufficiently senior divides the required work and then assigns it to others.

  - *Do you work alone when working on a project or as part of a team?*

The majority of interviewees responded that they work within teams of two or more people to complete projects. However, an interesting observation here is that working with others is not necessarily equated to *teamwork* with individual team members (or sub-groups) maintaining responsibility for their own component within a larger project team. (response: "*originally just myself and the technical director which did this sort of thing .. and then we'd work with the software guys .. they had their own software bits .. and the electronics guys worked on the electronics bit .. so it wasn't much of a teamwork thing …. it was like when I needed the software guy I would speak [to him]*")

- *Are projects divided up into the roles and capabilities of the specialised individuals?*

Recurring answer here was 'yes'.

- *How does communication occur between the different people/groups working on a project?*

There was no standardised communication structure identifiable here. From the interviews *distance* was a factor identified by the respondents which did impact communication. Those who worked in the same factory/space as the other team members found communication an easier process. Conversely, it was noted that software developers would often be contractors who would work remotely which made communication difficult.

- Are you given information about the whole of the project you are working on, or just the part of the project that you are specifically working on?

Those interviewed all indicated that they knew about the projects they were working on, even when they were only completing a small part of them. Of course, there is level of bounded knowledge here, in that there may be unknown unknowns, and/or the individual may be supplied with false information.

- *From the moment a project has been announced and the work divided, how is the work overseen and completed as the project progresses?*

There was no standardised oversight structure or set of processes identifiable here. It differed depending on a multitude of factors including the project itself, who the person was, and the size and structure of the employer. The following are examples of the variety displayed within the responses provided to this questions: (responses: "*it changed project to project*"; "*I do whatever [my boss] tells me to do until [] is happy with the finished product*"; "*the technical rep was supposed to be in total control of it or the senior vice president or his boss .. it depended what the projects were .. they kind*

of dished [the work] out to either the vice president or the technical director .. and then there's the engineering manager as well who was more on the building-it side rather than the physics-side"; "*It's my company .. I am the oversight*").

- *What happens when or if a previously unrecognised or unanticipated problem becomes apparent when completing a project?*

The common response was to notify those higher in the *chain of command* and any other groups that may be affected. If the problem was one which would ultimately affect the ability of the designer to produce the product asked for by the client, then senior staff interviewed said they would contact the client.

- *Are employees encouraged to come forward with a better solution?*

The responses here were particularly interesting. Those who were less senior indicated that they felt they could pass ideas on to their supervisors, but that the opportunity to do so did not arise very often as they were often given quite specific tasks to complete and they tended not to question these. Those in more senior positions indicated that they were always encouraging employees to come up with better solutions. However, this assertion was subject to caveat or a narrow interpretation of what this actually entailed. For example, one respondent stated they allowed employees to provide feedback, but that any subsequent decision making remained top-down and rigid; hence *feedback* did not equate to *action*. Another respondent welcomed new ideas but the example they provided of such an action was of minimal impact; (response: "*we always look for our employees to come up with new solutions, [and] ideas. Only the other week one of our junior engineers came up with a new way for using a bracket*"). The question of whether they would welcome more expansive, impactful feedback was not explored.

- *How detailed are the specifications you are given at the start of a project?*

The recurring response is that this depends on the client. An expansion of this from one interviewee was that if the client has worked with the company in the past and has built up a relationship of trust and understanding then in turn their specifications tend to be less detailed.

- *How binding are these? Do they allow for 'wriggle-room' or are the parameters always stringently defined and adhered to?*

The recurring answer is this was very much context dependent. On a spectrum responses here ranged from 'not being afforded the option' (response: "*I wasn't involved in [setting] the actual physical design at all*"), to 'it depends' (response: "depends what it is .. we always work towards [set national/international] standards .. we know we want [the product] to perform to 'this' level of standard .. that was on the majority of projects"), to 'a lot of freedom to improvise' (response: "*a lot of the things I was working on were already existing products .. so it was 'make this better'*").

- *What happens if you realise there is a better way of doing something which the client may not have realised but will require the specs be changed?*

For those interviewees working in direct contact with clients, a decision such as this very much depended on their relationship to the client. If it was an established relationship built up over time then they indicated they would be more willing to have such a conversation with them.

- *How much influence/flexibility do you have on projects you are given?*

Once a project was agree to with a client, the focus was very much on producing the agreed product to the agreed specifications. Within the designer, the work was divided accordingly. There was little opportunity to explore avenues of interest.

This could be contrasted with those working on academic research projects. The emphasis here was on producing what they had agreed to secure funding. However, greater willingness or perceived freedom was reported here to explore different

avenues to achieve the intended goal, though there were still limits (response: "*I didn't have the time to do my own thing that much .. I mean once they had given me a project to work on then I could be a bit flexible .. I could look at it from different aspects .. but again I couldn't really say 'oh this is quite an interesting thing .. why don't we look at this*").

- *How important is the cost of a ST to clients?*

This was repeatedly identified as *the most important factor* for any ST project (response: "cost is everything").

<u>Thoughts and Opinions of any Future Design Tools:</u>

- *Are you aware of any tool that can be used by the designers of security technologies for anticipating and mitigating negative social reactions to your designs?*

All respondents answered 'no' to this question.

- *What tools do you already use to identify social issues when beginning a project?*

All respondents answered 'none' to this question.

- *Do you think there needs to be different tools specifically for group-use and others specifically for use by individuals?*

Respondents were unsure as to whether or not different group-user versus individual-user tools needed to be produced. This lack of certainty is not surprising given that all respondents in the previous question noted that they did not use such tools. Hence the ability to base opinions to this question in past experiences is likely to be severely limited here.

- *When would be the best time in the design process to use the proposed tools?*

All interviewees indicated that any tool should be designed to be used at the beginning of the ST design process. This reasoning's behind this view included; (a) that it will prevent the need to modify the design later in the design process (response: "*it's got to be at the beginning .. because these things will affect the design .. and you don't want to go back .. so we've designed this and we are starting to build it, and now we have to change the design*"), and (b) that it will assist in the creation of clear design goals which can then be met.

- *Is it already too late once you're telling the individual engineers what to build?*

This line of questioning was to determine at what *level* in the design process a tool should be used. The general consensus was that this should occur at higher levels of management. This answer then led into the next question.

- *But given it's the lower-down people who actually have to build the ST, aren't they better placed to actually build something in a different way?*

The overriding view of those more junior interviewees was that they would prefer not to have the responsibility of implementing this tool. They would much rather be presented with a set of specifications *already modified by any tool* which they could then set about producing. (response: "*I would rather be presented with a set of specs already incorporating the social issues rather than being presented with a set of possible specs and potential social issues and then have to build it from there*")

- *How long should it take to learn how to use a tool for the first time?*

There was general resistance to any tool which takes a long time to learn. This appeared to be based on either *bad* experiences by the interviewees of previous courses which were viewed as a waste of time, by being too long and/or containing superfluous information.

- *Regarding the instructions for the tools, how long could they be before you stopped reading/following them?*

Again, there was an overriding desire for brevity here (response: "*literally a case of read for more than an hour ... I would be bored*"; "*if it's a tutorial then maybe half a day tops*").

- *Would it be valuable if they contain concrete examples of how each aspect of the tools are applied to a real-life or hypothetical design project?*

Consensus view was that the inclusion of examples was preferred.

- *Would you want to have somebody step you through how the tools are used, especially if you had never used them before?*

There was no consensus on this issue. This appeared to reflect the different approaches of the interviewees. Some were in favour of *immersion* (response: "*I think I would prefer just jumping in, with examples of how to use it*"). Others liked the idea of being shown how to use something by a trainer. While others indicated that they would like to work it out for themselves but that a *help line* they could contact if they needed would be valuable.

- *How long should any tool take to use?*

Regardless of the seniority of interviewee the consistent view was that the time taken to use any tool should be minimal. No interviewees gave a use-time of longer than one day. Times varied from one hour through to one day. Some of the respondents did add the contextual caveat of 'it depends on the project', indicating that it might be possible for a tool to take longer than one day to use and still be acceptable if the circumstances so dictated.

- *Is it realistic to think this time would be available if you could show the value of doing it?*

There was consensus that time would be made available if the tool was considered a valuable addition to the design process. The implication of this qualification is that if the value of a tool was not acknowledged then time might not be assigned to its use.

- *What format would you want the tool to be?*

There was no consensus on this issue. Different respondents preferred either digital or paper-based formats. Comments such as "*I very much like a point-and-click type thing*", "*an app would be good*" and "*I like things I can hold on to [so paper-based]*" were all recorded. The differences in opinions here can be framed as both beneficial and frustrating. They are beneficial in that they do not rule out any particular format of tool, hence maximising the available options when designing such tools. However, they are frustrating in that they do not provide much guidance on form, and with the absence of a majority opinion and tool designed in a single format may struggle to maximise its uptake.

- *Would a tool need to be bespoke to a particular security product (such as CCTV) or could their design be generalised so as to apply to all products?*

Given the responses here, it is possible to conclude the view that a balance needed to be struck between tools that are of general application and those designed for specific security products. There was a general acceptance that certain aspects of a tool may need to be bespoke depending on the ST at hand (for example it was noted that a CCTV camera is very different from a metal detector in terms of the technology involved). On the other hand there was acceptance that it would be impractical to create a bespoke tool for every single ST being designed.

- *Should it offer potential solutions based on previous controversies from other similar or related technologies?*

Most of the respondents here valued the idea of information on previous solutions in the form of suggested solutions.

- *Do you think your final products would be improved by the use of such tools, and that engineers in general would benefit from such tools?*

There was general consensus that final products may be improved by such tools but without expansion on exactly how they would necessary be better.

3.4 Discussion of Results

The Stage 2 interview questions were divided into three themes (education, working practices, and tool design); the significant results of which were presented in Chapter 3.2.1 above. Within this section the relevance and implications of these results are discussed, with the primary conclusions highlighted.

Education:

Table 3.2 provided a breakdown of the professional qualifications possessed by the interviewees. This included mathematicians, physicists, IT specialists, and a range of engineering specialities. While diverse, the education of all these individuals falls within a grouping which has become known internationally as STEM (i.e. science, technology, engineering and mathematics).

When interviewing the subjects on their individual STEM degrees, three common factors emerged:

1. ***Each interviewee's course comprised a set of core modules on top of which they would choose electives to assist in specialisation and/or the exploration of fields of interest.***

Thus if topics such as ethics and/or social impact are to be taught to STEM students they would have maximal impact if incorporated into core modules rather than electives; purely because all students within a degree will take the core modules.

2. ***There were no social-impact or ethics subjects offered; either within the core or elective subjects.***

This raises a number of related issues. Firstly it calls into question the level of impact that the rise of engineering ethics is actually exerting; especially if it is not filtering down into the education programmes of the current cadre of engineering graduates. Secondly, the obvious implication of point 2 is that the interviewees received no formal social-impact or ethics training as part of their STEM education.

On the specific issue of ethics, there were responses by some of the interviewed engineers that ethics did form a component of their engineering education in that engineers are taught to build bridges that don't fall down, and are built to standard. I question this characterisation of ethics; one which conflates competence as ethics. It is arguable that an engineer who builds a bridge that is stable and meets all relevant standards is merely an engineer competent at their job. To be an ethical engineer, (following on from the discussion of responsibilities within ethical engineering in Chapter 1.5.3) requires something more. The engineer must be prepared to challenge the bridge itself; i.e., its positioning, its impact on the environment, the materials used, its dimensions, whether it should be built, etc., rather than just building a stable bridge[116].

It should also be noted here that even if engineering ethics actually has an impact on the design of ST, by placing an onus on engineers to modify their work practices and products, many of those involved in the design of STs (mathematicians, physicists, etc.) are not engineers. No movement comparable to ethical engineering was identified through interviews with those from other STEM fields (i.e. mathematical ethics, physics ethics, etc.). Thus any overall impact of ethical engineering on the design of STs is further diluted.

3. ***There was a lack of conviction amongst the interviewees that they would opt for a subject on the social impact or ethics of their chosen fields should it be offered.***

Few of those interviewed expressed interest in taking social-centric optional subjects. There was an expressed concern over the relevance of such subjects. There was also

[116] On the ethics bridge construction see Winner (1980) for a fascinating discussion of the incorporation of racism into the design of the New York parkways, whereby social-class bias and racial prejudice was (literally) built into their construction. By restricting the height of the overpasses, busses (frequented by African-Americans and poor people) were prevented from using them, leaving them free for the automobiles of the white upper- and middle-classes.

the view that as there was already so much to learn within their chosen fields, any optional subjects would be better spent on their core discipline and/or specialisation. Finally there was simply a relative lack of enjoyment or interest in social-impact or ethics; at least when compared to mathematics, engineering, etc.

Working Practices:

Of the three interview themes, the questions relating to work practices elicited the greatest diversity of responses. This is understandable given the multitude of influencing factors operating on each interviewee in relation to this theme. The size of their companies, the nature of their employment (i.e., academia, industry, government, independent contractors), established working practices, whether they work as individuals or as part of larger teams, their experience, and the nature of the STs they have worked on will all influence the interviewees' opinions on how they perceive working practices. It was possible however to identify three common factors:

4. ***There exists a distinct lack of uniformity or standardised set of processes over the structure and conduct of ST design projects.***

How work on projects is divided, who works on them, how communication occurs within these projects, and how oversight occurs within a ST design project all differed. This was evident both between different sectors (i.e. universities as opposed to commercial industry), within a sector (i.e. different companies within the commercial sector), and even within a single company (i.e. the processes associated with how a ST moves from concept to prototype can differ within the same company from project to project).

5. ***The ability of an individual to directly influence the design of a ST appears to decrease the more junior the position of that individual within their employment structure.***

From the interviews it became clear that contact between a company and a client occurred at a higher level within a firm. As a rule junior staff were not involved in these negotiations/interactions. Within academia hierarchical structures also existed within laboratories and research projects; whereby the principal researcher who had been awarded the funding would direct the actions of more junior researchers and students.

Given the funding constraints within both research grants and a commercial budget it is understandable junior staff responded that they were not in a position to develop work of their own within a ST design project beyond that which had been pre-assigned to them from their managers.

There were, however, avenues for feedback. Junior staff responded that they would inform a more senior manager if they identified a problem when attempting to complete a project. And senior managers responded that they actively sought feedback from their junior staff when it came to solutions and ideas. However it was clear that such ideas/solutions could not be implemented independently by the junior staff, but would require senior approval.

Also while senior staff did say they encouraged employees to come forward with better solutions to problems, only concrete example of this happening could be provided by the interviewees. And the limited nature of this example ("*one of our junior engineers came up with a new way for using a bracket*") does call into question the level of influence junior staff can exert.

6. The overwhelming importance of costs to both clients and ST companies.

Money (in the form of expenses, sales and/or profits) is the single identifiable factor which ultimately determines both whether a ST design project is undertaken and the final design of that ST. It determines whether or not a company will bid for a government tender, it sets boundaries to research projects, and it limits the design possibilities. No sector (commercial industry, government, or academia) is immune.

A secondary factor which may influence how much a client is willing to spend is *trust*. Some responses indicated that if a relationship of trust already exists between the client and the ST developer then the client may be more amenable to altering their budgetary requirements. However from responses by senior staff, during a recession even a trusted relationship may have little sway over the amount of money a client is willing to spend.

<u>Tool Design:</u>

The results here were potentially the most informative of the three interview themes. The responses to questions on the structure of any future design tool were particularly

relevant given that the interviewees were the anticipated end-users of any future design tool.

7. ***There was no awareness of, or experience using, tools for assisting the designers/developers of STs in mitigating negative social reactions to their products.***

The importance of this point is that future design tools cannot rely upon a presumed level of experience held by end-users acquired through using similar tools. An additional consequence is that the creation of any future design tool may have to begin from a blank page if there are not existing tools from which to draw experience.

8. ***Any design tool should be intended for use upstream in the design process before any construction of a prototype/product begins.***

This was the unanimous view of all interviewees, and will influence the design of any future design tools.

9. ***The time taken to: i) learn how to use a tool, and ii) to actually use it, should be minimised. In any event use-time should not exceed one day.***

Extended training courses were generally derided by the interviewees as a waste of time. However there was no consensus over the preferred design of any training (i.e., whether it was by trainers, online tutorials, a manual, etc.).

What was agreed upon was that a design tool would be employed if the time taken to use it could be self-justified as a worthwhile expense of resources. To this end a period of one day was considered the most that would be spent on such a tool.

10. ***There was no consensus over the format of any future design tool.***

The views here ran the gamut of purely paper-based, through to computer programmes and apps.

11. ***A balance needs to be achieved between creating a tool intended for general application and one which is sufficiently bespoke to a specified technology.***

It was recognised that any design tool must have a certain level of general application otherwise a new tool would need to be created for each technology/design-project.

However, the view was also shared that there were benefits to tailoring a tool such that it focussed on specific STs.

12. ***There was support for the concept of creating design tools to assist designers/developers of STs in mitigating negative social reactions to their products.***

This was good to see as it provides at least a level of external justification for the aims of this research project.

4. Assessment Criteria Identified From Combining Stages 1 & 2

As depicted in Figure 1.1 *Macro-level structure for conducting the research project*, Stage 1 of this thesis involved the analysis of a series of case-studies of controversial STs whilst Stage 2 involved interviews with engineers and scientists involved in the production of these technologies. Both stages are discussed in detail within chapters 2 and 3 respectively.

The results of these two Stages are of fundamental importance to this research project. By combining the information derived therein it is possible to develop *assessment criteria* by which to judge any methodological tool seeking to identify and mitigate negative social reactions to security technologies during their design process.

This chapter sets out these *assessment criteria* for utilisation within Stage 3 of the project whereby existing tools/methods are assessed against them. The examination of existing tools/methods and their subsequent assessment against these criteria will be undertaken within Chapter 5 below.

4.1 The Assessment Criteria

By combining the results of the twelve case-studies into controversial security technologies and the results of interviews with the engineers and scientists engaged in the production of these technologies, a list of eleven criteria are produced below. Each one is accompanied by a short describing sentence and a brief discussion by way of explanation and justification.

1. For use before building commences: Any design tool should be intended for use before the physical act of constructing any prototype/initial-ST commences; in other words, at the stage when design requirements and design specifications are being identified and agreed upon.

One of the strongest messages to come out of the initial interviews was the unanimous opinion that any design tool should be applied as far upstream within the design

process as possible. Certainly that it should be applied before any physical construction (or coding for digital products) of prototypes or products begins.

While this view has been adopted as *Criterion 1* here, there are challenges with this approach which need to be acknowledged. These challenges, and the counter-arguments against them, are discussed below.

Firstly *Criterion 1* implies that the specifications of the ST will remain steady throughout the period of its design which may be a number of years, and this is not always the case. During the design period the technology will remain at the mercy of external factors and influences; including changes in - public opinion, the security threat, and the political landscape, amongst others. For example with the UK National Identify Card[117] there was a significant drop in public support for this technology during the lifecycle of its construction; from a reported '79% for - 13% against' around 2002/3 (Home Office 2003) to '47% for - 51% against' by 2006 (ICM 2006). And the UK National Identity Register was physically destroyed before national rollout following a change in government (Mathieson 2011).

While this concern is valid it remains a fact that STs will *always be* susceptible to external forces during their design process. The real question therefore becomes whether it is more efficient to; (a) address all social concerns after the first version of the ST is developed, or (b) to address identified concerns both at the beginning of the design process and then reassess the final product against the prevailing social climate at the end? Given that it was the dominant view of those interviewed in Stage 2 that it is much easier to make changes to a design before work on a project has been undertaken, addressing as many social concerns as possible at the beginning of a design project appears the best strategy in that; (i) it does not prevent later social assessment and (ii) it will hopefully minimise the scale of any secondary assessment. Also given that the commonalities of controversy identified in Stage 1 represented causes of social concern that *repeatedly* arose across multiple STs, it would be illogical to ignore such concerns when the opportunity to address them for the minimal amount of effort presents itself at the beginning of a project.

[117] See Chapter 2.5

This last point (on the repeating nature of social controversies) also addresses a second concern with *Criterion 1*; namely whether it is feasible to identify potential sources of social unacceptability within the design of a particular ST *before* that technology is released into the public domain? The commonalities of controversy identified in Chapter 2 demonstrate that this is possible, though with two caveats:

1. Newly arising and/or previously unidentified commonalities will not be consciously addressed;
2. Unanticipated social reactions to STs are to be expected given you are introducing a (potentially complex) ST into an intrinsically complex and diverse system that is a society[118].

The consequence of this discussion being; while it may be optimal to attempt to minimise negative social responses to STs upstream in the design process (and hence the creation of *Requirement 1*), guaranteeing a positive response is nevertheless impossible.

2. Intended for senior-level use: Senior designers/developers should be involved in using any design tool, so as to make full use of their personal experience and authority to ensure any design decisions based on the tool are afforded sufficient weight by all those involved in the design process.

While the interviews indicated the presence of feedback opportunities for staff within ST companies, the presence of strong hierarchical structures was also identified. Junior staff interviewed indicated that when work was assigned to them within a project their overriding objective was to complete that work as directed. This top-down authority structure was acknowledged by senior managers. Rather than complaining about this structure, those junior staff interviewed preferred being assigned work where the design specifications had already been set as it helped direct their work.

[118] One can visualise the complexity of society and its interaction with STs as an incredibly complicated engineering endeavour; one where accidents are to be expected. In his book *Normal Accidents* Perrow (1984) discusses complex systems where accidents are expected events given that the complexity of the system leads to *complex interactions* (i.e., unfamiliar, unexpected, or unplanned sequences of events) and *tight coupling* (i.e., the absence of a buffer or give between two items – what happens to one happens to the other). By extending this analogy to include the society in which the technology will operate, one can accept that negative societal responses to STs should be anticipated.

Additionally the task of deciding whether a tendered ST was feasible fell onto senior management within the commercial firms included in the interviews. These senior managers were engineers and scientists with considerable experience in developing STs and in directing teams and workflows.

Given this existing structure and division of responsibilities it would be prudent to employ design tools at senior levels. This would allow the results of such to infuse the entire project as work is subsequently divided and filtered down into the wider workforce.

3. Cannot involve external actors: Any design tool must be able to work under the condition that no public engagement/involvement whatsoever is allowed before or during the design process.

As discussed in Chapters 1.3 & 1.4, secrecy is often a defining characteristic of STs; be it secrecy to promote a commercial advantage or secrecy imposed by a government. This secrecy may be in relation to specific capabilities of a ST that the public are aware of (such as the detection rates of various substances by airport whole body-scanners[119]) through to secrecy over the very existence of a ST (such as the US PRISM programme brought to light by Edward Snowden (Greenwald et al 2013).

Public involvement or engagement during the initial design of these technologies was non-existent. Indeed for whole-body scanners in the UK, the Department for Transport (DfT) produced an Interim Code of Practice for their use in 2010 without public involvement (see DfT 2010c) and after many years of trials in UK airports. A public consultation was not held until after publication of this Interim Code when the DfT solicited information on a revised (non-interim) code. Notably this public consultation was only about the *future operation* of these scanners, not about their *legitimacy or on-going presence* in UK airports (DfT 2012b).

Criterion 3 is recognition of the need to create design tools which operate within the existing paradigm governing the design of STs. It is not an affirmation of this paradigm, nor does it make claims as to the efficacy or legitimacy of the current systems.

[119] See Chapter 2.2

4. Produces design specifications, not policies: The design tool must produce output that is usable by the designers. Most likely this will be design requirements and design specifications which will influence the STs they are about to produce.

The purpose of any design tool must be to assist the designers themselves. At a fundamental level it is foreseeable this could be achieved through the identification of potential sources of social controversy within the proposed ST such that the designers can then consciously address them. To possibly add extra value it may prove useful if the design tool could also flag previously enacted technical solutions to the same or similar problems arising in previous STs.

This task of producing design specifications is separate to that of governments or agencies in developing policies regarding the use of these STs. While of undoubted importance, such activities fall outside the scope of responsibilities that can reasonably be attributed to those scientists and engineers engaged in designing these technologies.

5. Output must add value: The output produced by the design tool must be such that it enhances the final products, thereby adding value to them.

Criterion 5 is similar to *Criterion 4* in that they are both focussing on the output of any future design tool. While *Criterion 4* focusses on the type/form of output produced, *Criterion 5* focusses on what this output must achieve; namely the 'adding of value'.

Value here is a relative term as what is considered *valuable* will differ depending upon the views, motivations, and responsibilities of the individual beholder. The design company may value the profits derived from the sale of a ST. A law enforcement end-user may value the way a ST enhances their ability to prevent or detect criminal activities. A false-positive citizen targeted incorrectly by a ST may value a society whereby the provision of security is not allowed to negate competing rights.

Regardless, these different conceptualisations of *value* are not necessarily incompatible. A ST which is acceptable to the citizen will not become the target of social resistance, thus will be more likely to enjoy continued application. If it *works*

then it can be endorsed and promoted by end-users. The result being the manufacturer of a socially acceptable STs enhances their opportunity to maximise their sales and thus providing the greatest return for their initial investment.

6. Usable in any workplace: The tool must be usable regardless of the design environment or how this environment is organised.

STs are designed in a variety of environments. The interviewees in Stage 2 included individuals working to develop STs within private companies, universities, governments, and as private contractors. As a result any design tool produced to assist these individuals should not be created with a single work-space in mind if it is to enjoy the widest possible application. Nor should any tool assume the existence of a standardised template for how the work within a particular organisation (or even a particular project) is divided and undertaken. Unless a tool is being created specifically for an organisation (or a particular project) then a level of flexibility and generic application should be assumed and built-in.

7. No prerequisite expertise required: The ability to effectively use this design tool should not be premised on the user possessing a particular qualification or formal social-science education/training beyond any training-course/instruction-manual which accompanies the tool itself.

There was a complete absence of formalised social science training from the STEM educated interviewees in Stage 2. There was also a less-than-enthusiastic response by some of the interviewees to the possibility of taking social science focussed courses/subjects at university as part of their STEM degrees. On the assumption that this scenario is repeated within the wider ST-design workforce, it would appear counterproductive to require the end-users to change so as to make a design tool (supposedly created for their benefit) more usable. The preferred alternative is to create a design tool which requires no specialised skills or prior formal qualification/education on the part of the end-user. This criterion does not preclude

specific training on the use of the tool itself in the form of a training course and/or an instruction manual.

8. Minimal use-time required: The amount of time required to use any design tool must be made as short as possible.

The interviewees in Stage 2 supported the concept of design tool to assist the developers of STs in anticipating and mitigating negative societal responses to their predations. However, they were also clear that there was a limit to the amount of time which could/should be assigned to the use of such tools. Few interviewees provided specific timeframes, but those that did ranged from a few hours to one day. Given these comments *Criterion 8* seeks to ensure that any design tool strives to minimise the amount of time required to be devoted to its use.

9. Addresses multiple, diverse controversies: Any tool must be able to address all of the commonalities of controversy identified within Chapter 2 of this thesis.

Stage 1 of this research project identified 43 commonalities of controversy arising from previous STs. It would be wholly unrealistic to create a separate, bespoke design tool for each of these commonalities and expect the developers to actually use them at the start of every design project. As such *Criterion 9* dictates that any design tool must be capable of dealing with multiple commonalities.

10. Adaptable to all STs: Any tool must be applicable (and/or if necessary, adaptable) to any ST regardless of form or function.

Similar to *Criterion 9*, given the incredible diversity of STs, any design tool created must be able to be applied to any ST as it would be near impossible to create a completely bespoke tool for every existing and potential technology. This would also impose additional training burdens on the developers. That being said, it was a view from some of the interviewees that the ability to modify a particular design tool so as to

ensure its relevance to the ST being developed would be a useful capability. The idea proposed was for the tool to comprise of a central core applicable to all technologies with additional modules which could be attached or removed so as to tailor the design tool to specific categories of ST.

11. Not based exclusively on 'yes'/'no' answers: Given the complex contextual nature of the identified commonalities of controversy, a tool based exclusively on binary 'yes'/'no' responses will be inappropriate.

The case study analysis in Stage 1 highlighted the contextual nature of many of the identified commonalities of controversy. This characteristic brings into question the effectiveness of designing a tool *based solely* on binary, definitive responses (such as yes/no, acceptable/unacceptable, etc.) when those responses are determined by the context within which a ST is designed and deployed.

Such binary approaches may still be valuable and find a place within any design tool as there were certain commonalities which presented themselves as definitive choices[120]. However as these were the minority, *Criterion 11* cautions against adopting this approach as the rule rather than the exception.

[120] See Chapter 2.14.1 (specifically point 5) above for examples of this.

5. Stage 3 – Literature Review of Existing Tools/Methodologies

Stage 3 of this project involves assessing existing candidate models, approaches, and/or assessment tools with two goals in mind. The first goal is to determine if (based on both the steps involved in carrying them out and the nature of output produced) any of these candidates can sufficiently meet the eleven criteria as set out in Chapter 4.1 above without further modification? If they do then such candidates could operate as *off-the-shelf design tools* for assisting the developers of STs in identifying and mitigating sources of social resistance within the upstream design of these technologies. If no candidate can achieve this end-result then the second goal is to identify what (if any) value could be extracted from these candidates for incorporation into new design tools capable of meeting the criteria from 4.1 above.

The methodology employed to achieve these goals is expanded upon below in Chapter 5.1 with the results presented in Chapter 5.2. This is followed by a discussion addressing the implications of these results for the future direction of this research project in Chapter 5.3.

5.1 Methodology

Meeting the two goals outlined in Chapter 5 above requires an assessment of existing candidate models, approaches, and assessment tools based on the assessment criteria produced in Chapter 4.1. The information pertaining to each candidate which needs to be extracted is primarily the steps-for-use and the nature of the output. This data is to be collected through the use of *multiple case studies*. A general discussion of this methodology has already been undertaken in Chapter 2.1 above.

The scope of the case studies to be conducted and the time required to complete them are both shorter than for those in Chapter 2.1, in that the particular phenomenon's of interest (i.e., *steps-for-use* and *nature of the output*) are usually more settled and thus easier to extract. Again as per Chapter 2.1 the data sources are

documentation based; including journal articles, books, and instruction documents related to the use of these candidates.

An alternative methodology for obtaining this information would be interviews with end-users of these candidates. I have rejected this approach as the primary sources are both available and inherently stable, which are important characteristics given that I am interested in 'typical versions' of these candidates. There is always the probability interviews with end-users will elicit information on how different end-users modify these candidates to suit their various needs, and by not conducting interviews I will be excluding such information. However, this level of detailed information is not required to complete the aims of Stage 3 and hence is not undertaken herein.

5.1.1 Method

There were three steps involved in conducting the research within this Chapter. These are described below.

Step 1. ***Decide on candidate tools/methodologies:*** Any existing decision-making or assessment tool/methodology/model is a potential candidate as a design tool for meeting the aim of assisting the designers of STs. While this freedom opens the door to a range of diverse options regarding the shape of any future design tool, from the practical perspective of conducting this research project it creates a massive challenge. Given the existing constraints of time, word-length, and available resources, it is impossible to conduct a bespoke case-study for every individual candidate within this dissertation. To address this problem while still maximising the number of potential candidates included, three actions have been adopted herein:

i. *To combine related models/methodologies possessing largely generic traits into single case-studies wherever possible and appropriate* - For example the single case-study *Checklists* (see Chapter 5.2.1) provides a generic series of steps fundamental to any checklist-centric model.
ii. *To wherever possible include tools which have been designed with STs in mind as potential targets for use* – For example, Bruce Schneier's *five-step process* for

analysing and evaluating security systems, technologies, and practices (see Chapter 5.2.6)

iii. *To triage* – Two of the essential elements governing how tools are designed within this research project are that; (a) these tools must produce useful information for the designers of STs, and (b) these tools must be able to operate in an environment of secrecy (i.e., without requiring the engagement of external actors). As such certain classes of existing assessment models were excluded from consideration, which included those that provide policy-focussed information and guidance, and those where public-engagement was a central or essential component[121].

I readily admit these approaches may both exclude some viable candidates from consideration and necessarily gloss over some of the granular differences between individual models when collated. Nevertheless I see them as necessary and suitable compromises allowing for the incorporation of as many candidates as possible within the practical constraints of this research project.

Step 2. ***Conduct focussed case-studies to identify the steps involved and outputs produced from the candidates identified in Step 1:*** Short case-studies were then carried out on the (categories of) candidates identified in Step 1. These are presented in Chapter 5.2 below. The specific focus of these case-studies is to identify both the

[121] As a direct result of this action a number of prominent models have been excluded from this analysis. These include the following non-comprehensive selection of examples. *Social Impact Assessments* (SIA); defined as "*the process of identifying the future consequences of a current or proposed action which are related to individuals, organisations and social macro-systems*" (Becker 1997, p.2). Underpinned by extensive public involvement (Becker 1997) SIA has been viewed separately as either a process which facilitates interest-group negotiations or as a means of assisting political decision makers (Barrow 1997). Also omitted are the various forms of *Technology Assessment* (TA); "*a scientific, interactive and communicative process which aims to contribute to the formation of public and political opinion on societal aspects of science and technology*" (Butschi et al 2004, p.14). The purpose of TA is to assist policy makers by offering knowledge and advice to those struggling with the decisions that will shape the technology in question (Butschi et al 2004). By association this omission also covers *participatory Technology Assessment* (pTA); "*the class of methods and procedures of assessing socio-technological issues that actively involve various kinds of social actors as assessors and discussants*" (Joss and Bellucci 2002). As well as *Constructive Technology Assessment* (CTA); a core idea of which is that developers of technologies should enter into discussions with a diverse range of concerned parties and actors during the design process of a new technology, thus contributing to its future successful development (Schot and Rip 1996).
In an attempt to counteract the effects of shutting off the public from the ST design process, my design is acting as a form of *proxy* or *virtual society* with which the engineers can engage.

steps involved in undertaking the particular methodology/model, as well as the nature of the resulting output. Given this narrow focus, these case-studies are not (and are not intended to be) comprehensive accounts of the candidates. Topics including the history and motivation behind their design are not included, and neither are critiques of their respective strengths and weaknesses based on previous use.

Step 3. ***Compare the results of Step 2 against the eleven assessment criteria identified in Chapter 4:*** The results of Step 2 are then examined against the assessment criteria produced in Chapter 4.

5.2 Results

Table 5.1 below lists the case-studies of models and methodologies examined within Chapters 5.2.1 through 5.2.6. The results of this process were then employed in Appendix *J* to assess the likelihood that these candidates could be utilised as design tools within the remit of this research project with minimal or no modification.

Table 5.1 List of models and methodologies for assessment

	Candidates
Ch. 5.2.1	**Checklists**
Ch. 5.2.2	**Impact Assessments** – privacy impact assessments – surveillance impact assessments – ethical impact assessment framework for information technologies – anticipatory technology ethics
Ch. 5.2.3	**Frameworks** – basic frameworks o Friedman's framework for the governance of information security o Da Veiga and Eloff's framework for the governance of information security – applied frameworks o dual-use decision framework for technology governance

Ch. 5.2.4	**Design-Focussed Approaches** – value sensitive design – privacy by design
Ch. 5.2.5	**Quantitative Assessments** – cost-benefit analysis – multi-criteria decision making
Ch. 5.2.6	**Miscellaneous Tests with a ST Focus** – Bruce Schneier's five-step process – ACLU's necessary and defensible test

5.2.1 Checklists

Without meaning to overstate the obvious, at its most fundamental level a checklist is a formalised series of questions/activities that the end-user is required to undertake/assess before making the requisite 'mark' indicating the answer to the question and/or the completion of the activity. It is easy to disregard checklists as being little more than simple *box-ticking exercises* however they have been extensively and successfully used in many fields, including aviation (to prevent incidents during take-off, flight, and landing), medicine (to improve surgery safety, reduce infections, assist in diagnosing, etc.), as well as engineering, policing and legal practice (Gawande 2011; WHO 2008; Thomassen et al 2010). According to Gawande they can provide protection against both "*the fallibility of human memory and attention, especially when it comes to mundane, routine matters that are easily overlooked under the strain of more pressing events*", and people "*lull[ing] themselves into skipping steps even when they remember them*" (2011 p:36). They also allow for the minimisation of errors within incredibly complex engineering endeavours; such as the construction of high-rise buildings involving the coordinated efforts of multiple specialist professions.

<u>Steps involved</u>

1. Create a checklist from scratch, acquire a pre-designed checklist, or modify an existing checklist so it is bespoke to your needs/organisational-structure/etc.
2. Work through all of the questions/statements/activities comprising the checklist in the order that they are presented by marking the checklist items in the manner required.

Output

The output of a checklist is often a binary yes/no response to whatever the question asked, though may include other information (for example, current flight speed and heading, or the number of forceps accounted for at the beginning and end of an operation).

5.2.2 Impact assessments

Based on the focus of this research project, *impact assessments* (IA) provide a means by which to assess a proposed technology or programme by focussing on a specified element of its impact upon a specified population (for example the privacy impact of a particular design of whole-body scanner on the total population of airline passengers and/or specific sub-sets of this population - such as women, ethnic minorities, different religious groups, children, etc.). The various forms of *impact assessment* (IA) incorporated here include *surveillance*, *privacy*, and *ethical*.

Surveillance IA which is a very recent concept is essentially based on the privacy IA model[122], hence it is easy to produce a combined summary of these two forms of assessment and their output. Ethical IA's however, are more diverse in their construction, hence two different models are outlined below:

1. Ethical impact assessment framework of information technologies
2. Anticipatory technology ethics

While the *ethical impact assessment framework* is accompanied by a fairly clear series of steps for using this framework, *anticipatory technology ethics* is somewhat vaguer in its description of use. However, it is still possible to discern a prescribed methodology.

(privacy impact assessments; surveillance impact assessments)

Steps involved

The following set of steps represents those required for completing an IA; either *privacy* or *surveillance* (see Wright and Wadhwa 2012; Wright and Raab 2012, p.615):

[122] It differs by looking beyond just privacy-impact by also examining social, financial, political, legal, ethical and psychological issues. It also requires consultations with a wider range of stakeholders (Wright and Raab 2012).

1. Determine whether an IA is necessary
2. Identify the IA team, setting its terms of reference, resources and time frame
3. Prepare an IA plan
4. Determine the budget
5. Describe the proposed project to be assessed
6. Identify stakeholders
7. Analyse the information flows and other impacts
8. Consult with stakeholders
9. Determine whether the project complies with legislation
10. Identify risks and possible solutions
11. Formulate recommendations
12. Prepare and publish the report
13. Implement the recommendations
14. Ensure a third-party review and/or audit of the IA
15. Update the IA if there are changes in the project
16. Embed privacy awareness throughout the organisation and ensure accountability

Output

As per Step 12, the physical output of an impact assessment is a report on the object, project, service or programme under consideration. However, from Wright and Wadhwa (2012) it can be inferred that the true value of an IA lies in the *process* of undertaking the impact assessment and the knowledge gathered as a result; with the report assisting in documenting this process. Output could therefore also be taken to include any resulting remedial actions or the mitigation of issues resulting from this process.

(ethical impact assessment framework for information technologies)

Developed by David Wight[123] this model begins with a framework structured around five major ethical principles pertaining to the design and operation of information technologies, namely; respect for autonomy, nonmaleficence, beneficence, justice, and privacy and data protection. Each of these principles is further broken down into a

[123] See Wright (2011)

number of values and/or issues[124]. For each of these value/issues (and for the majority of the principles) questions have been posited to draw out possible sources of unethical design within the information technology to which it is being applied. This process is intended to occur as follows (Wright 2011):

Steps involved

1. Based on the framework of principle, values, and issues, direct the accompanying questions to the technology developers and/or policy-makers "*to facilitate a consideration of the ethical issues which may arise in their undertaking*" (p.204).
2. Use these principles, values, and issues to engage with, and generate debate amongst, stakeholders. Tools which may facilitate this process include; consultations and surveys, expert workshops, checklists of questions, an ethical matrix, ethical Delphi, consensus conferences, and citizen panels.
3. To complement and assist in the answering of the questions posed in Step 1 (beyond engaging and consulting with stakeholders as per Step 2), a number of additional practices and procedures have been highlighted. These include; use of risk assessments, determining accountability within a project third-party reviews and audits, and providing information and responding to complaints.

Output

The adequate examination of ethical issues arising within new technologies, services, projects, policies or programmes by stakeholders to allow for necessary mitigating measures to be undertaken before deployment (Wright 2011).

(anticipatory technology ethics)

Developed by Philip Brey, anticipatory technology ethics (ATE) "*distinguishes three levels of ethical analysis: the technology, artefact and application level*" (2011, p.18), which each involve different processes of analysis. At the *technology* level, morally relevant features of both the wider technology and its subclasses are studied. The *artefact* level refers to specific black-boxes created from these wider technologies; for example x-ray body scanners are artefacts of nuclear technology. The *application* level

[124] See Appendix *G* for a complete breakdown of these principles, values, and issues.

refers to how an artefact is used and the contexts within which it is used (Brey 2011).

Steps involved

The series of discernible steps for conducting ATE divides this process into three stages; *forecasting stage*, *identification* (ethical analysis) *stage*, and *evaluation stage*.

1. (forecasting) Engineers provide ethicists with an understanding of the future developments of the technology at the *technology* level as described above.
2. (forecasting) For both the *artefact* and *application* levels forecasting and technology assessment (TA) models/tools should be utilised to provide ethicists within information on the likely future emergence of *artefacts* and *applications*.
3. (forecasting) Additionally for both *artefacts* and *applications* expert-surveys and round-tables should be conducted incorporating a range of actors, including; engineers, technology forecasters, TA experts, historians and sociologists of technology, and marketing experts.
4. (identification) According to Brey "*[a]t this stage, descriptions of the technology [as acquired in the forecasting stage] are cross-referenced with ethical values and principles*" (2011, p.23). This is achieved by employing an ethical checklist[125]. The technology is applied to the values and principles contained within to determine how it frustrates and/or enables these values within real-word conditions.
5. (identification) Survey technology ethics literature to identify additional ethical issues not contained within the ethical checklist and use bottom-up ethical analysis to evaluate the *artefacts* and *applications* against these.
6. (evaluation) The importance of the identified ethical issues is assessed to determine how likely they will become significant societal issues, and how they may conflict with other rights/values[126].

Output

Results of the *evaluation* stage can be used to guide how a technology is developed. Moral responsibilities can be assigned to relevant actors in relation to the *artefact* and *application* levels so as to ensure ethical outcomes. Recommendations can be produced for policy-makers (Brey 2011).

[125] See Appendix *H* for Brey's checklist.

[126] Brey has provided no guidance on how this evaluation step is to be successfully undertaken.

5.2.3 Frameworks

While the term *framework* has taken on numerous meanings within different contexts, the conceptualisation being applied here is that of: *a supporting or underlying conceptual structure of interlinked items intended to act as a support or guide for:*

1. *the achieving of a specific objective*
2. *planning or deciding something*
   and/or
3. *the creation of something that expands this structure into something practical*.[127]

There are three fundamental points arising out of this conceptualisation. The first is that frameworks can (and do) take on many different shapes and forms; there is not a single 'correct' construction. Secondly, given the versatility of frameworks (useful for many tasks and not restricted to a single discipline) there is a plethora of examples in the marketplace and academic texts.

Thirdly and perhaps most importantly, is that while a framework can help guide the end-user as they seek to address a problem, it will not by itself simply provide them the solution they seek. The end-user will still need to undertake other tasks to achieve their objective. The framework may inform the end-user as to what information must be considered or tasks undertaken when making their decision, but it will not *prima facie* provide them either the information itself or the answer being sought. It provides them with the basic skeleton onto which other methodological elements may be superimposed to create a dynamic tool or model, without determining for the end-user what these other elements should be.

To expand these points, I have included three examples of frameworks with possible application to security technologies. The first two, under the heading *basic frameworks*, represent diverse examples of skeletal frameworks which have not been built upon. They provide a visual representation of information to be considered by the

[127] This multi-faceted definition represents the compilation of a variety of different dictionary sources, including; Oxford, Cambridge, www.businessdictionary.com, and whatis.techtarget.com.

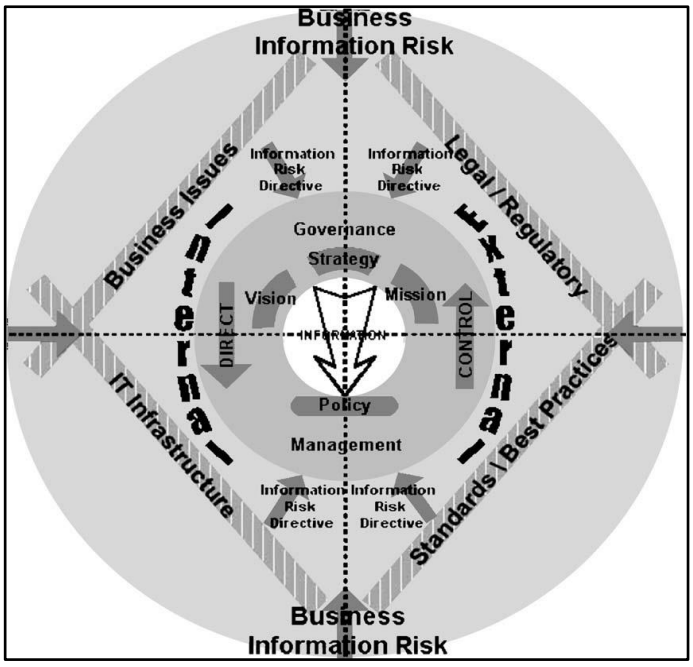
end-user but no guidance of how they should make use of this information. As such there are no Steps involved or Output subsections included here.

This is followed by the *Dual-use Decision Framework for Technology Governance*, an example of what I have termed an *expanded framework*; referring to a framework which has undergone expansion to incorporate additional processes. There extra processes allow for the addition of meaningful Steps involved and Output subsections.

(Basic frameworks)

Figure 5.1 is a framework for the governance of information security whose aim is to integrate information security into the corporate governance structure of an organisation (Posthumus and Solms 2004).

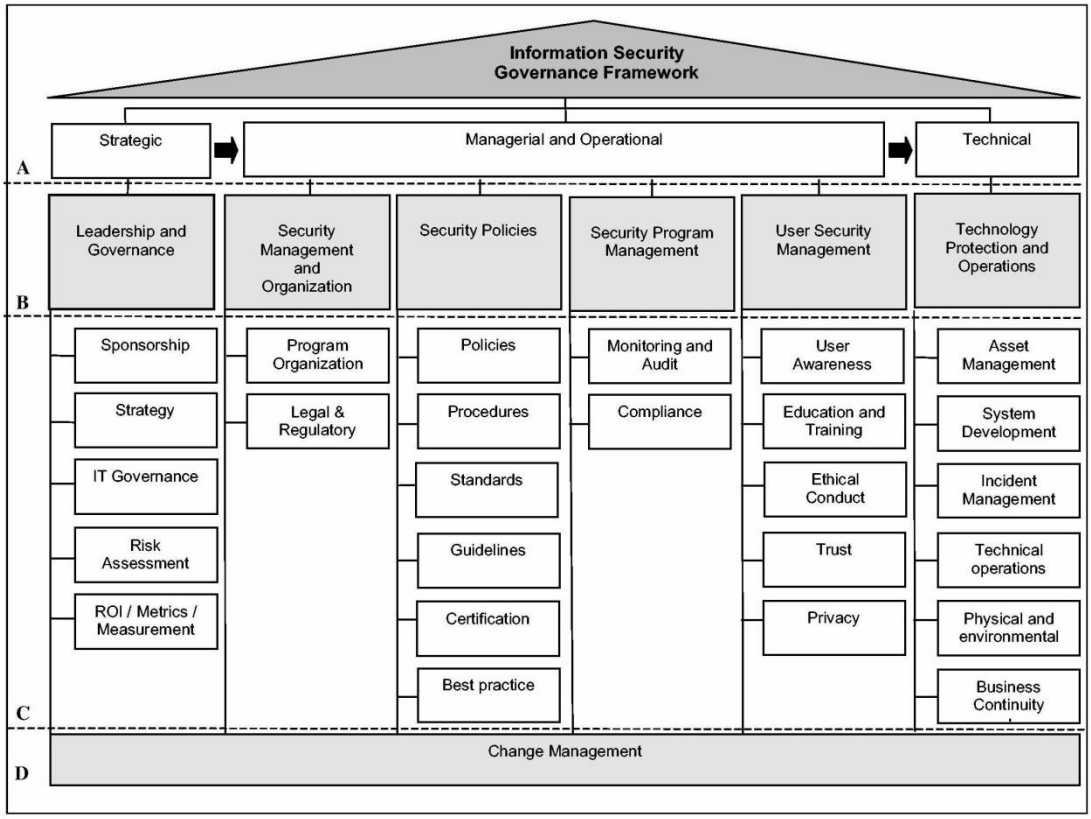
Figure 5.1 Framework for the governance of information security



While this framework contains information as to actors, artefacts, and activities, as well as incorporating the use of arrows to imply dynamic interactions between entities, it is not intuitive how this framework is to be applied.

Figure 5.2 also represents a framework for the governance of information security; one that encompasses technology, processes and people. It is intended to be utilised to cultivate an information security culture within an organisation so as to minimise risks to information assets (Da Veiga and Eloff 2007).

Figure 5.2 Information security governance framework



The presentation of this framework is possibly more intuitive to risk managers within an organisation than the previous example for, according to Da Veiga and Eloff (2007), it is based upon a number of established information security frameworks, including ISO 17799, ISO 27001, and PROTECT. Nevertheless the creators have not provided instructions for how their own governance framework should be utilised by end-users. The approach adopted in Figures 5.1 and 5.2 (of presenting just the bare framework with no accompanying information on methods of utilisation) can be contrasted against the following example.

(Expanded frameworks : *Dual-use decision framework for technology governance*[128])

This model is a "*decision framework that policy makers can use to assess the risk that individual emerging technologies will be misused for hostile purposes, and to develop tailored governance strategies*" (Tucker 2012, p.67). There are three interconnected processes comprising this decision framework; those being: (i) *technology monitoring* for detecting candidate emerging dual-use innovations; (ii) *technology assessment* to

[128] See Appendix *I* for the complete Decision Framework.

determine the likelihood of misuse and the potential for bespoke regulation; and (iii) *governance measures*.

Steps involved

The steps involved in using this decision framework are as follows:

1. Monitor technological developments to identify emerging technologies in the biological and chemical fields with the potential for misuse.
2. Assess their risk of misuse using four parameters of; accessibility, ease of misuse, magnitude of potential harm, and international diffusion.
3. If aggregate misuse risk from Step 2 is adjudged low: no urgent governance needed, just maintain monitoring.
4. If aggregate misuse risk is medium to high: assess governability of the technology using five parameters of; embodiment, maturity, convergence, rate of advance, and international diffusion.
5. If aggregate governability of the technology from Step 4 is adjudged low: focus on informal governance measures.
6. If aggregate governability is medium: consider also adding soft-law governance measures.
7. If aggregate governability is high: consider all measures (informal, soft law, and hard law)
8. If risk of misuse appears exceptionally grave and imminent: consider governance measures beyond those normally adopted under this framework
9. Use a cost-benefit analysis to produce a package of governance measures for reducing misuse of the technology for an acceptable cost, and in a manner acceptable to all major stakeholders.

Output

The Dual-Use Framework identifies and prioritises those emerging dual-use technologies which are likely to warrant governance measures, as well as guiding on the most effective types and combinations of such measures (Tucker 2012).

5.2.4 Design-focussed approaches

These are approaches aimed at influencing the development of a technology by incorporating certain values into the research and design processes. This approach of incorporating societal values during the design-stage is as much an ethos for the construction of technological artefacts as it is a range of practical models. As such it is not always possible to reduce these approaches down to a series of basic steps. While this has been achieved below for one model (*value sensitive design*), it was not possible for the second approach examined (*privacy by design*). As such for privacy by design in place of Steps involved I have included its seven foundational principles. These constitute a universal reference framework that may be developed into detailed criteria for application (Cavoukian 2011a).

(value sensitive design)

Emerging in the mid-1990s, value sensitive design is a principled approach to the design of a technology what seeks to comprehensively account for human values throughout the design process. It comprises a tripartite methodology incorporating investigations of a conceptual, empirical, and technical nature (Friedman 2004; Friedman et al 2008). A basic set of processes is as follows:

Steps involved[129]

1. (conceptual) Begin by identifying the core aspect(s)[130] central to the current project. Use this/these as the starting point to draw out the others.
2. (conceptual) Systematically identify the key direct and indirect stakeholders (i.e., respectively those who interact directly with the technology or its output, and those impacted by the system).
3. (conceptual) Systematically identify the harms and benefits for each group identified in Step 2.
4. (conceptual) Based on the harms and benefits for the stakeholders as identified in Steps 2 & 3, now identify the corresponding key values which are engaged.

[129] Based on Friedman et al (2008).
[130] The three core aspects to choose from here are; value, technology, or context of use (Friedman et al 2008).

5. (conceptual) Conduct a conceptual investigation for each identified key value to accurately define those values as they appear within the project.
6. (conceptual) Examine potential conflicts for the key values (e.g. privacy versus security, etc.).
7. (empirical) Interview stakeholders to elicit information about the values.
8. (technical) Based on the identified values and value-conflicts, make explicit possible design trade-offs and their effects.

Output

The output here includes; the knowledge gained during the entire process which will be valuable to the organisation undertaking it, and the identification of different design options and (hopefully) will translate into a more value-sensitive product.

(privacy by design)

Developed and driven by Ann Cavoukian[131] during the 1990s, the philosophy of privacy by design evolved from efforts to incorporate *fair information principles* directly into the design, operation, and management of information communication technologies. The results of which became known as *privacy enhancing technologies* (PETs) (Cavoukian 2009b; Cavoukian 2013). As discussed above, as privacy by design is foremost a concept rather than a prescriptive model, I have included here the seven foundational principles upon which it resides which themselves can be used as a framework to create applicable tools.

7 Foundational Principles[132]

1. *Proactive not reactive; Preventative not remedial*: the aim is to prevent privacy infractions occurring rather than waiting for privacy risks to materialise.
2. *Privacy as the Default Setting*: Personal data should be automatically protected within any IT system or organisation by default.
3. *Privacy Embedded into Design*: Privacy is built into the original design as an essential core component and not bolted on afterwards.

[131] Information and Privacy Commissioner, Ontario, Canada.
[132] Based upon Cavoukian (2011b)

4. *Full Functionality – Positive-Sum, not Zero-Sum*: Privacy by Design seeks to accommodate all legitimate interests and objectives rather than promoting false dichotomies (e.g. privacy *versus* security is rejected in favour of privacy *and* security).
5. *End-to-End Security – Full Lifecycle Protection*: Because Privacy by Design is built-in from the beginning, personal data is protected to ensure privacy from the moment it is collected until deletion.
6. *Visibility and Transparency – Keep it Open*: All stakeholders (including users and providers) should be able to determine that all privacy promises/objectives are being met by the relevant business-practice/technology.
7. *Respect for User Privacy – Keep it User-Centric*: The interests of individuals should be prioritised through privacy defaults, notices, and user-friendly options.

Output

According to Cavoukian (2013) there is an emphasis on practical results within privacy by design. This includes such goals as; enhanced accountability and trust in relation to technologies and organisations, transparent data processing, enhanced accountability, preventing unauthorised access to and/or processing of data, etc.

5.2.5 Quantitative assessments

By *quantitative assessments* I am referring to those methodologies where the intrinsic processes involve the assignment of numerical values, and resulting numerical outputs are used to either influence or determine any subsequent decision-making. Two possible candidates are outlined below, these being; *cost-benefit analysis* and *multi-value decision making*.

(cost-benefit analysis)

A cost-benefit analysis is a methodology for decision-making based on whether the benefits of an action outweigh its costs. It can be utilised to answer questions such as; when faced with a number of options (*A*, *B*, *C*, etc.) which of these should be selected based on the respective costs of each? (Quah and Toh 2012; Mishan 1976). While there are many different mathematical methods of varying complexity that can be

employed when conducting a cost-benefit analysis, essentially the entire process reduces down to the following three steps.

Steps involved[133]

1. Decide which items are relevant to include (and exclude) on both the *cost* and *benefit* sides of any analysis.
2. Compute the values of the items included in Step 1.
3. Compare the cost and benefit values thereby providing informed advice to assist the decision-maker.

Output

The outputs are the numerical values produced within the cost-benefit analysis. However just like any other form of evidence, it remains the prerogative of the decision-maker as to the weight they choose to assign this output when exercising their duty to arrive at any final decision. This seemingly objective process is complicated by the fact that the things the decision-maker values (and how much they value them) may be different from what others groups within the same society value.

(multi-criteria decision making)

Multi-criteria Decision Making (MCDM) refers to the process of deciding on the optimal decision out of a set of alternatives in an environment of multiple, often conflicting, criteria. It is often used to model practical problems such as the purchase of a house or choosing a suitable office-space.

There exist a number of different MCDM processes. The one outlined below has its basis in those proposed by Roy (1990) and Balton and Stewart (2002).

Steps Involved[134]

1. *Problem structuring*: develop a thorough understanding of the problem and define the objectives.
2. *Identification of potential alternatives*: identify alternatives based on combinations of independent choices.

[133] See Quah and Toh (2012).
[134] These steps were produced with assistance from Dr Sonia Toubaline, and are based on an unpublished work by her.

3. *Construction of a family of criteria*: this is the criteria used for basing your decision on. It can be a time consuming process and may require complicated calculations, surveys, expert advice, etc.
4. *Selection of the problematic*: 'problematic' refers to methods used to present the alternatives – for example; rank ordering the alternatives, sorting them based on predefined categories (i.e., accepted, rejected), etc.
5. *Evaluation of the alternatives and problem parameters*: evaluate each alternative against each criterion.
6. *Multi-criteria analysis*: aggregate the selected criteria to produce the optimal solution.

Output

The outputs are the values produced by the MCDM approach and the subsequent ranking of alternatives. However, just like any other form of decision-making it remains the responsibility of the person(s) tasked with making the decision whether or not to apply the MCDM output.

5.2.6 Miscellaneous tests with a ST focus

(*Bruce Schneier's five-step process*)

Steps involved

Security expert Bruce Schneier has devised a five-step process for "*analyz[ing] and evaluat[ing] security systems, technologies, and practices*" (2006, p.14). At a bare minimum this process requires contextual analysis of the proposed ST (including how, where, and why it will be used), risk analyses, and cost-benefit analyses. Step 2 may require the engagement of security experts and/or those possessing specific information on possible attacks. Additionally, while not specifically stated, public engagement may be needed to successfully address Step 5 and possibly Steps 4 and 1.

1. Determine what the assets are you are trying to protect. Answering this question requires developing an understanding of the scope of the problem faced.

2. Determine what the risks are to these assets. Questions to be answered here include; what exactly is to be defended?, from who?, how and why will it be attacked?, and what are the consequences of a successful attack?
3. Determine how well the security solution mitigates the risks identified in Step 2. This is to ensure that the security solution *actually solves* the security problem. To achieve this determination requires an examination of how the security solution interacts with its environment (both its successes and its failures).
4. Determine what other risks the security solution causes. All security solutions may result in unintended consequences in the form of new security problems. The goal is to ensure that these new problems are smaller than the old ones.
5. Determine what costs and trade-offs are imposed by the security solution. All security systems impose costs and require trade-offs; be it money, resources, time, convenience, freedoms, etc. It is essential to understand these trade-offs.

Output

According to Schneier (2006 p.15):

> These five steps don't lead to an answer, but rather provide the mechanism to evaluate a proposed answer. They lead to another question: Is the security solution worth it? In other words, is the benefit of mitigating the risks (Step 3) worth the additional risks (Step 4) plus the other trade-offs (Step 5)? It is not enough for a security measure to be effective … we need to do the things that make the most sense, that are the most effective use of our security dollar.

(ACLU's necessary and defensible test)

Created in the wake of 9/11 and the subsequent clamour for new security measures and powers, this analytical tool was devised by the American Civil Liberties Union to test whether a specified privacy-infringing anti-terrorism measure is capable of achieving its stated purpose (Roy 2005). It requires the following three questions be addressed but without providing prescriptive guidance on how best to achieve this.

Steps involved[135]

1. Does the government already possess the resources to combat the problem that the new proposal is meant to address?

[135] See Roy (2005 pp.51-52).

2. Is the proposal narrowly tailored so as to limit the adverse impact on civil liberties?
3. Does the proposal genuinely combat terrorism, or does it represent a wider legislative change unrelated to September 11?

Output

Ideally the output would be in the form of an answer to the question; will a proposed security measure achieve its stated purpose? This is presented as the logical and mandatory prerequisite requirement for any proposed anti-terrorism measure (Roy 2005).

5.2.7 Comparison against assessment criteria

The case-studies undertaken in Chapter 5.2 on the methods and outputs from existing models and methodologies are compared against the eleven assessment criteria outlined in Chapter 4. The results of this comparison are presented in Appendix *J* through a series of tables, each representing a single model/methodology. In the following section (Chapter 5.3) the overall implications of this evaluation process for the future direction of this research project are discussed.

5.3 Discussion of results

By compiling the information obtained through the comparisons of the different candidate models/approaches against the eleven assessment criteria, as presented in Appendix *J*, the following conclusions are drawn.

a) None of the candidates can be classified as perfect *off-the-shelf* design tools for assisting the developers of STs in identifying and mitigating social resistance upstream in the design process while operating under an environment of secrecy. When compared against the assessment criteria outlined in Chapter 4.1 all of the candidates experience difficulty meeting at least one of these criteria.
b) As a caveat to (a) above, while all of the candidates would need some modification to meet the previously identified assessment criteria, the changes required vary

enormously between the different candidates. This ranges from minor adjustments (see the Ethical Impact Assessment Framework for Information Technologies) through to the complete reconstruction of a method (see Privacy and Surveillance Impact Assessments).

c) There is value to be drawn from many of the candidates. Even if a particular candidate proves inappropriate, often elements of it can be adopted within a completely new design tool.

Based on these conclusions, in Table 5.14 below the candidates are categorised according to their ability to add value to future design tools.

Table 5.2 Potential of candidates to add value to future design tools

Could act as the foundational basis for a future model	
Expanded Frameworks	Could be designed from scratch to operate as a future design tool.
Ethical Impact Assessment Framework for Information Technologies	This combination of a framework with associated questions could be modified and expanded to encompass all commonalities and STs.
Can provide elements for (or operate as an element for incorporation into) a future model	
Checklists	Can operate as an element of a broader model.
Anticipatory Technology Ethics	Provides ethical checklists, and expands the role of engineers within the design process.
Basic Frameworks	Can provide the skeleton upon which to develop a comprehensive model
Value Sensitive Design	The weighing of harms and benefits and the identification of key values.
Cost-Benefit Analysis	Can operate as an element of a broader model.
Multi-Criteria Decision Making	Can operate as an element of a broader model.
Not suitable for use within this project	

Privacy & Surveillance Impact Assessments
Privacy By Design
Bruce Schneier's Five-Step Process
ACLU's Necessary and Defensible Test

From this analysis *Expanded Frameworks* (of which the *Ethical Impact Assessment Framework for Information Technologies* constitutes an example) appears to be the most viable identified candidate for forming the basis of a design tool that could meet the aims of this research project. It is worth reiterating here that this conclusion, and the concomitant rejection of the other candidates, is based purely on the capacity of these candidates to meet the assessment criteria produced in Chapter 4; assessment criteria formed from combining the interviews of STEM practitioners engaged in designing STs with the case-study analysis of previous controversial security technologies. No other conclusions or assumptions should be drawn from this rejection. These models/approaches are immensely valuable tools, they are simply inappropriate for meeting the designated requirements of this research project within the relatively narrow field that is the design and development STs in the environmental constraints that currently exist.

5.3.1 Implications for the research project

The implications of these results for the direction of this research project were:

a) The identified need to create new bespoke design tools for assisting the developers of STs in identifying and mitigating sources of possible social resistance within their technologies upstream in the design process.
b) That these tools should be based on an extended framework model.

While these implications have provided the future direction for this research project (i.e. the need to create new design tools based around expanded frameworks) they do not provided guidance as to how such an undertaking should be completed; i.e., they tell us what to build but not how to build it. This problem is addressed at the start of

Chapter 6 which begins by providing rules for the construction of future design tools[136]. These rules are then applied within the construction of two such tools.

[136] See Chapter 6.1.

6. Stage 4 – Creation of Bespoke Tools

By this stage in the research project, forty-three commonalities had been identified from the analysis of the case studies into controversial security technologies. These were subsequently categorised into seven designated categories (see Chapter 2.2). Next it was concluded that while these was much value to be found within existing technology assessment tools, none were appropriate 'off-the-shelf' vehicles for simultaneously addressing all of these commonalities. Especially when designing a security technology within the specific secrecy constraints which dominate this industry. As a result of this limitation it became necessary to create new tools to meet the goal of assisting the designers of security technologies in anticipating and mitigating possible negative social reactions to their future products upstream in their design processes.

Design rules governing the creation of these new tools, gleaned from the work undertaken in the prior chapters, are detailed in Chapter 6.1 below. These rules are then applied to create two design tools based on the design requirements outlined in Chapter 4 and in compliance with the rules outlined in Chapter 6.1. These are:

1. Framework of Common Controversies within Security Technologies
2. Designing for Socially Acceptable Security Technologies

Both of these are presented in Chapter 6.2 below. This chapter details why they have taken on the forms they have, as well as an explanation of how I envision these tools being used.

6.1 Rules governing the design of the bespoke tools

The *multi-criteria conundrum* faced when producing design tools within this project is to find a way to creating tools which can simultaneously address *all* of the forty-three identified commonalities of controversy *while at the same time* being applicable to *all* security technologies. To assist in overcoming this challenge (and taking account of both the results of the individual case studies in Chapter 2 and the responses of the developers of security technologies in Chapter 3) there are five identifiable and

essential design rules which can be used to dictate the form of any relevant tool produced within this project. I believe any tool must achieve *all* of these five rules if it is to stand a meaningful chance of adding value to the design process of security technologies, regardless of the technology or the nature of the social controversy it may evoke.

Five Essential Design Rules

1. *Any tool should be applicable to all security technologies as they are defined within this project*[137]. While at first glance it is tempting to limit the scope of any tool to some arbitrary subset of security technologies (for example; digital, biometric, aviation, weapons, autonomous, surveillance, CCTV, etc.) I have rejected this approach for two reasons. Firstly the identification of forty-three commonalities of controversy which had common application across a range of ostensibly disparate technologies gave weight to the argument that tools should be developed based on the central underlying purpose of security technologies (i.e., the provision of security as conceptualised within the definition of security technologies) as opposed to subsets of these technologies based on physical form or intended uses. Secondly, the melding together of different security technologies within the same platform and/or the encouragement of function creep within both the design and the operation of security technologies undermines attempts to *black-box* these technologies by form or capability. This in turn undermines any attempts to create tools bespoke to a particular technology type.
2. *Any tool must be able to simultaneously address all forty-three commonalities identified within Chapter 2*. As stated throughout, the goal of this project is to create tools for assisting the designers and developers of security technologies in anticipating, and thereby avoiding, potential design choices which are more likely to lead to social resistance upstream in the design process. To achieve this goal these tools must be able to address all of the identified commonalities for the following individual and grouped arguments. Firstly, having identified forty-three commonalities it would be both unrealistic and counterproductive to create

[137] This being; the product of an engineering endeavour which seeks to deter, prevent or detect crimes, and/or enhance the security of individuals, their property, or the state (including its infrastructure). This may include potentially lethal technologies, but does not include technologies restricted to military use.

individual, separate tools bespoke to each commonality and realistically expect a developer/designer to apply them all. It is highly doubtful they would possess the time, resources, inclination, or the patience to systematically work through forty-three individual tools, all with ostensibly the same goal. However I believe a *single tool* which can successfully identify potential sources of social resistance for all forty-three commonalities is much more likely to see use[138]. The second grouping of reasons relates to the observed tendency for security technologies to evoke negative social reactions from both unobvious sources or as a result of unintended, unanticipated interactions/outcomes. Given how difficult it is to predict *ex ante* the source and nature of all controversies arising from either the operation of a newly designed security technology or the novel application of an established security technology, any tool designed within this project should at least possess the capacity to address all forty-three identified commonalities so as to maximise its potential positive impact. Thirdly designers should avoid arbitrarily restricting any examination they undertake of future social controversies given the harm each individual commonality can cause; regardless of how rarely that commonality arises. In the earlier case-study analysis (Chapter 2) not all of the identified commonalities appeared as often as the others. For some of these commonalities, their over or under representation within this project will be a direct consequence of the small sample size of technologies examined herein. However for others, such as those falling under Category 1: *Physical or Mental Harm*, the instances of these commonalities arising will be generally lower given the nature of security technologies currently being produced; there are more new security technologies for the collection and manipulation of data being produced than security technologies with the capacity to directly physically harm individual targets. While the temptation exists to create tools which only address the most popular commonalities, this should be avoided as there is not necessarily a direct correlation between the regularity of a particular commonality arising and the

[138] This point is not intended to negate the value of tools which seek to address specific individual social issues which persistently arise as a source of controversy (such as privacy, dual-use, data processing, etc.). For issues such as these which are already in both the publics' and regulators' sights, and hence more likely to receive media attention, it is wholly understandable for designers to employ additional tools which focus individually focus on such issues so as to thoroughly examine any future potential negative impact.

damage this commonality arising will cause to a security technology. To illustrate, while security technologies possessing the capacity to cause unintended deaths[139], serious/permanent injuries, disfigurement, and/or loss of mobility/senses[140] (for example, mosquitos[141] and less lethal weapons including Tasers, baton-rounds[142], and possibly millimetre-wave active denial technologies[143]) comprise a tiny subset of the total range of available security technologies, the potential level of public backlash (both by scope and scale) is very high. Alternatively, even a rarely observed commonality possesses the potential to result in disproportionate levels of harm if it is arises within a ST that becomes very widely used.

3. *Any produced tool must be usable by the intended end-users*. The implications of this seemingly self-evident requirement are twofold. Firstly that the level of resources required for using a tool is not so high as to prevent or discourage their use. By *resources* I include both the time required to use the tool and any financial burden incurred. Secondly, that any tool should be designed such that it can be reasonably assumed that the intended end-users of the tool (i.e., the developers and designers of security technologies) will already possess the intellectual skills to do so; i.e., they should not have to acquire specialist social-centric academic qualifications to complete the tool, neither should the engagement of external experts be required. The justifications for this design requirement (that any produced tool must be usable by the intended end-users) are based on both; (a) the interviews with designers/developers of security technologies where they described the restricted nature of designing security technologies[144], and (b) a rejection of time-consuming, resource-laden approaches as unrealistic. Time constrains set by government tenders, the quickly evolving nature of cybercrimes, the speed at which technologies evolve, and the need to remain ahead of

[139] Commonality 1f: *Could the ST cause a fatality?* With both the Framework of Common Controversies within Security Technologies tool and the Designing for Socially Acceptable Security Technologies tool.
[140] Commonality 1e: *Could the ST cause physical injuries?* With both the Framework of Common Controversies within Security Technologies tool and the Designing for Socially Acceptable Security Technologies tool.
[141] See Chapter 2.12
[142] See Chapter 2.13
[143] See Global Security (2011)
[144] See Chapter 3.4

competitors within the security marketplace, all combine to deny designers the luxury of long lead-times before a design project can commence.

4. *The tool must be usable without any direct public interaction or input*. This is probably the most contentious design rule for a tool whose explicit purpose is to anticipate negative public responses, yet it is a necessary rule for reasons of secrecy[145]; a point confirmed by the interviews undertaken[146]. This design rule is not intended to devalue public engagement; it is merely a recognition of existing constrains governing the design of security technologies. Nor does this rule expressly forbid public involvement if a particular design project is not bound by secrecy. Indeed it may be possible to orchestrate the conduct of *some* design projects so as to include public involvement despite secrecy provisos[147]. However given the reality of constraints operating on many security technology design projects I believe any design tools should be constructed to operate successfully in the most restrictive of environments to thereby afford the tool the widest scope of usability.
5. *Any design tool must produce useable results*. What constitutes *useable results* in this circumstance is determined by the target audience; i.e., the designers and developers of security technologies[148]. The goal of any tool should be to produce practical design recommendations specific to the security technology under consideration. Given that the focus of this project is to influence the *upstream* design of a security technology *before* it is produced, the tools should emphasise the identification of possible sources of social controversy in such a manner which focusses the mind of the user towards identifying design alternatives which can address these sources for inclusion in the initial blueprint of the security technology. It is not the intention of this project to produce tools for critiquing completed security technologies; the focus is to remain firmly fixed on identifying upstream design opportunities.

[145] See Chapter 1.4 for discussions on national and commercial secrecy, and the implications of introducing secrecy into the conduct of a research and design project
[146] See Chapter 3.2.1
[147] See Chapter 8.4 for a discussion of potential methods whereby the public may be involved in security projects despite restriction to the contrary
[148] As discussed in Chapter 1.2

6.1.1 A critique of possible approaches for addressing the multi-criteria conundrum while maintaining a single design tool.

In an attempt to address the *multi-criteria conundrum*, and thereby identify suitable methodological structures[149] for the creation of design tools, I examine the structure of a number of existing methodologies. The resulting structural analysis is presented in Table 6.1 below followed by a discussion of the results[150].

The methodologies examined here are utilised for assessing security technologies/interventions, providing security, and/or reducing crime. For example, *Problem Orientated Policing* and *Situational Crime Prevention* and crime reduction measures utilised by police. Problem orientated policing moves police away from *reacting* to crimes towards analysing why specific crimes reoccur in particular places so as to develop pre-emptive solutions. Similarly situational crime prevention involves manipulating the environment in a permanent manner to reduce criminal opportunities (Newburn 2007). The *ISO27001 Framework for compliance* represents best-practice specifications that are used by businesses and organisations worldwide to develop their Information Security Management System[151]. *Multi-Criteria Analysis* and *Cost-Benefit Analysis* are decision-making tools used by private companies, law enforcement, and security agencies to assist in decision making and the allocation of resources. *Privacy Impact Assessments*, as the name implies, assess the possible impacts on privacy of policies, interventions, and technologies. They are used by agents of the state and by private companies, and may be either voluntary or a mandatory legal requirement (for example see Homeland Security (2006)). The remaining examples (*Five-Step Process*, *Dual Use Decision Framework*, and *Surveillance Impact Assessments*) are all relevant in different ways to STs and hence their inclusion here, however they are not employed to the same extent as the other methodologies.

[149] By structure I refer to both the step-by-step processes which comprise a methodology as well as the nature and form of the produced results

[150] This examination also utilises some of the data collected in the Chapter 5 case-study assessments of existing methodologies

[151] For more information see www.itgovernance.co.uk/iso27001.aspx#2

Table 6.1 Structural Assessment of Existing Methodologies

Style of assessment tool	Examples	Flexibility of Process	Flexibility of Output	Consistency of Process	Consistency of Results
A structured set of repetitive questions and/or processes	▪ Five-Step Process for Analysing and Evaluating Security Systems ▪ Dual Use Decision Framework ▪ Privacy Impact Assessment / Surveillance Impact Assessment	Low→ Mid[152]	Mid	Mid→ High	Mid
A prescriptive methodology where the format of the results is pre-defined	▪ Multi-Criteria Analysis ▪ Cost-Benefit Analysis	Low→ Mid[153]	Low[154]	High	High
An open framework	▪ ISO27001 Framework	High	High	Low	Low
A malleable methodology where the processes and format of results are both adaptable.	▪ Problem Oriented Policing ▪ Situational Crime Prevention	Mid→ High	Mid→ High	Low	Low→ Mid

To expand upon the columns in Table 6.1 above:

***Style of assessment tool*:** Outlines four categories of existing assessment tools. I separated these tools on the basis of both the *processes employed* by which the relevant assessment tools arrive at some form of output *as well as* the *nature of the output* itself. These two factors were chosen as they directly impact upon the two criteria which form to comprise the *multi-criteria conundrum*; i.e., finding a way to create tools which can simultaneously address all of the forty-three identified *commonalities of controversy* while at the same time being applicable to all *security technologies*.

[152] Bad for different technologies and commonalities
[153] Can add/subtract different variables to the process
[154] Bad for different technologies and commonalities

***Flexibility of Process*:** 'Processes' refers to the series of steps or actions required by an assessment methodology so as to achieve its purported purpose. It is my belief that an assessment tool with rigid processes will not be able to simultaneously address a large number of disparate commonalities, and will likely prove inefficient in operation. To expand; if the processes are rigid in nature (i.e., low flexibility) and hence cannot be easily amended, then the ability to assess future technologies with hitherto unrealised capabilities is not assured if these new capabilities are not already encompassed by the existing process structure. Conversely, highly flexible processes which can be modified to encompass novel features on future technologies are more likely to add value to the upstream design process of future security technologies. Additionally the greater the flexibility of process, the more likely a single methodology will be able to address all of the forty-three commonalities of controversy identified within Chapter 2; thus avoiding a situation where the end-user is trying to artificially pigeonhole health and safety or financial factors into an assessment methodology designed to address privacy or dual-use issues. It must be noted here that I am not suggesting there exists a direct relationship between the *level* of flexibility of process and the *quality* of those processes. An assessment tool with an inflexible set of processes may still produce high quality results while one with highly flexible processes can equally produce poor quality results.

***Flexibility of Output*:** 'Output' refers to the manner and form by which results are presented. This may range from a numerical value (including amongst others a probability, a monetary amount, or the sum of multiple factors), the ticks/crosses of a checklist, through to purely qualitative output in the form of comments or actions to be taken. Where the methodology employed is so well established and the format of the output is so highly prescribed such that universal rules for a 'correct output' exists (such as the score from a multi-criteria analysis or the quantification of a cost-benefit analysis) the methodology will possess low flexibility of output. Conversely where the end-user has considerable discretion in how they choose to present the results, either due to the qualitative nature of results, a choice of acceptably correct output modes, or the absence of consensus over what constitutes correct output then the methodology in question can be interpreted as possessing high flexibility of output. It

is my contention that successful security technology assessment tools of the type envisioned within this project require relatively high output flexibility; otherwise they will not be able to simultaneously address disparate commonalities of controversy. A methodology with low flexibility-of-output that produces only a single output measure, especially one which is narrowly defined (such as 'probability of causing death per use'), will quickly become irrelevant when the commonality of controversy changes ('probability of death' may be relevant for less-lethal weapons but is irrelevant when dealing with *fair-use* principles and data-mining technologies).

Consistency of Process**:** A process which comprises the repetitive application of the same required steps (regardless of what is being targeted) will have high consistency-of-process. This will be a methodology where the steps to be undertaken are universally accepted and followed, or one where there are an accepted series of *core* steps with minor variations to ancillary steps for dealing with nuanced circumstances. High consistency-of-process does not necessitate low flexibility-of-process. A process can be flexible in its application to deal with different targets and circumstances while following consistent, accepted steps for each application.

Consistency of Results**:** A methodology will possess high *consistency-of-results* when two different people/parties applying the same methodology will arrive at the same result(s). For example, mathematical processes (i.e., statistical tests, construction of preference/utility curves etc.) if conducted using accepted methodologies and the same input data *will* consistently produce the same output (barring user error) and hence possess high consistency of results. Open-ended questions and/or questions without some criteria or methodology for determining the answer will possess low consistency of results.

6.1.2 Optimised approaches for addressing the multi-criteria conundrum while maintaining a single design tool.

From Chapter 6.1 above, the five identified requirements for any future design tool are:

1. It should be applicable to all security technologies;

2. It must be able to simultaneously address all forty-three identified commonalities;
3. It must be usable by the intended end-users;
4. It must be usable without direct public interaction or input;
5. It must produce useable results.

By applying these requirements to the structural analysis from Table 6.1, the following conclusions are reached:

Processes with high flexibility will be better able to deal with diverse security technologies and address different commonalities. The questions asked, data collected, and methods of investigation will be modifiable to meet these two requirements. Conversely processes with low flexibility will struggle here when the security technology contains features not specifically catered by the existing processes. And as discussed earlier, the diverse nature of the different commonalities reduces the possibility that an inflexible process will possess the intrinsic capacity to cater for all of these without the inbuilt ability to adapt.

It will be easier to tailor highly flexible processes and outputs to meet the needs of end-users given that these individuals are not a homogenous group. They possess different skill sets, work with different mediums, have different priorities, and work under differing conditions and restrictions. Rigidity here will restrict both usability and the propensity for use when the processes and/or outputs either *are not usable* by end-users, *does not add value* to the design process, or *does not recognisably add value* to the design process.

Similarly given the diversity of end-users; the ability to produce results that these different individuals/groups will find *usable* (and the subjectivity embedded within such a concept) requires highly flexible outputs. An output whose form and nature is meaningless is useless to the designer of a security technology.

Regarding the necessary absence of public engagement, the level of flexibility of process or output should not directly affect the achievability of this requirement.

It is concluded therefore that any design tool will need to score *highly* in *both* flexibility of process *and* flexibility of output. As such from Table 6.1 those assessment tools which comprise either categories 1 (a structured set of repetitive questions and/or

processes) and 2 (a prescriptive methodology where the format of the results is pre-defined) should be ruled out given their relatively low flexibility of both process and output.

In comparison, categories 3 (an open framework) and 4 (an adaptable methodology where the processes and format of results are both adaptable) both represent potential candidate methodologies for design tools given their relatively high flexibility of process and output.

As a word of caution, it should be acknowledged that the relatively low levels of *consistency of process* and *consistency of results* within categories 3 and 4 may have a negative effect on the quality of both the processes and outputs of these two categories. This will not necessarily be the case, however it does open the door to arbitrary output and processes which lack rigour. Replication of results here is also not assured[155].

That said, if we do not adopt a consequentialist mentality, any lack of consistency of processes or results only becomes framed as a *problem* if replication is treated as a goal or aim of the design tools. As an alternative, if we promote such concepts as *thinking outside the box*, *re-framing a problem*, or *encouraging creative ethical thinking* as our goals, then replication becomes less of an issue as each situation is deemed unique.

6.2 The tools

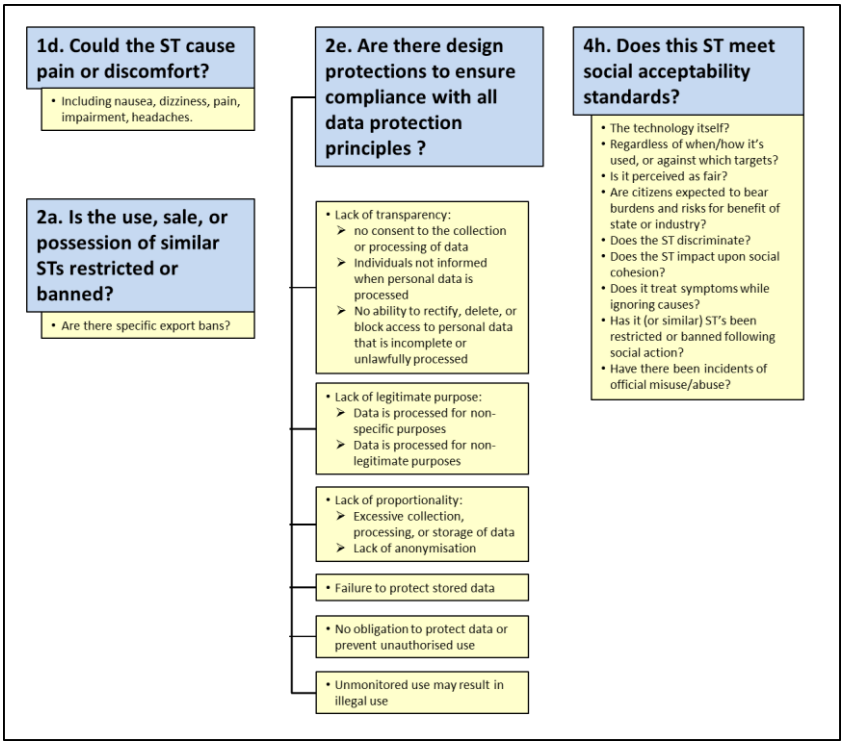
6.2.1 Framework for Common Controversies within Security Technologies

The *Framework for Common Controversies within Security Technologies* (FCC) presented in Appendix *K* represents a category 3 - *open framework* style of design tool as described in Table 6.1 above. Adopting the form of a stand-alone framework it is the less prescriptive of the two tools developed, while still providing an identifiable structure to assist with use.

[155] These are problems not unique to the design tools proposed within this research project; they also affect the other established assessment tools cited as examples of category 3 and 4 methodologies within Table 6.1

The FCC reproduces the forty-three commonalities identified in Chapter 2 divided into their seven categories. Additionally, each of these commonalities is accompanied by information in the form of questions, caveats, statements and/or clarifications. This accompanying information was distilled from the case-studies used to identify the commonalities and represents some of the more granular information that might otherwise have been lost by just presenting the commonalities themselves. The amount of additional information presented in the FCC varies considerably depending on the specific commonality; a point illustrated within Figure 6.1.

Figure 6.1 Extracts from FCC



Here commonalities 1d and 2a are quite specifically focussed and consequently very brief; 1d is accompanied by a list of symptoms which have constituted pain or discomfort in the past, whilst 2a is accompanied by a single question pertaining to export bans. By way of contrast commonalities 2e and 4h have broader remits; a fact reflected by the volume of additional information accompanying this commonalities. 2e focusses on compliance with data protection principles (an area with many sub-components), while the social acceptability standards addressed in 4h have proven diverse in nature.

Also the ability to discern the intrinsic nature of a commonality based on a plain reading varies considerably. Using the selection from Figure 6.1, while 1d '*Could the ST cause pain or discomfort?*' is relatively self-explanatory, the meaning of 4h '*Does the ST meet social acceptability standards?*' is not apparent from its wording. As a result, for commonalities such as 4h more accompanying secondary information may be necessary.

Some of the identified benefits of adopting this framework-based approach are that frameworks provide a foundational structure from which to begin any assessment processes. Additionally, being a basic framework existing at the highest level of abstraction there are no formalised instructions or rules governing how the FCC is to be used; these will need to be decided upon and produced by the end-user. While this process requires an initial expenditure of effort, ultimately the final product may better reflect the capabilities, resources, and organisational requirements/restrictions of the end-user than could be achieved if they followed a set of pre-constructed generic rules.

Paradoxically, these strengths can also be interpreted as weaknesses of the framework design. A framework that is provided to guide investigations may become counter-productive if the end-user applies it prescriptively such that they refuse to investigate (potentially valid) issues deemed to fall outside the scope of the framework itself. Additionally, if the end-user produces poor rules for applying a framework, any value it may have potentially produced could be negated.

On the FCC specifically, I anticipate its strengths as:

a) It provides a clear visual representation of the diverse range of social-acceptability problems which have impacted upon the operation of security technologies in the past.
b) It alerts the end-users to problems which they may not yet have anticipated.
c) Its application is not restricted by either commonality or technology.
d) It provides a basic frame onto which different end-users can attach secondary processes that facilitate design decision-making tailored to (i) meet the specific

needs of the designer and (ii) sympathetic to factors bespoke to the security technology being designed.

Specific weaknesses of the FCC include the following:

a) In isolation it is an abstract framework, hence does not provide guidance to the designer without the attachment of further processes. This fact reopens the door to the multi-criteria conundrum and the issues identified within Table 6.1; i.e. the secondary processes will need to be able to accommodate all possible security technologies and all identified commonalities of controversy.
b) The creation of secondary processes increases the work burden on the end-user before they can usefully use the FCC.
c) Best practices regarding the optimal secondary process for each commonality has not yet been determined. This will only come through using the FCC and applying different secondary processes so as to develop a corpus of knowledge on optimisation.

Moving beyond how the FCC is used is the question of what constitutes *successful* use. This is a question without a definitive answer given that it can be legitimately approached from both objective and subjective perspectives; see Chapter 7.1 below for an extended discussion on the challenge to measuring success. To summarise this discussion here; *objectively* one could define and measure success as the absence of social controversy and/or resistance to those future security technologies which are designed using the tools suggested herein. The difficultly here is that when success is defined as the absence of a phenomenon occurring it is incredibly difficult to determine conclusively whether this absence was the result of changes made because of the design tool or from some other factor(s); the classic *correlation versus causality* problem. *Subjectively*, success is determined by how different individuals and groups choose to frame such a result. This may include outcomes including; the identification of previously unrecognised problems, new design requirements, the modification or removal of existing design requirements, the uncontroversial adoption of a new security technology, or even either the abandoning of a proposed security technology or the absence of any changes.

6.2.2 Designing for Socially Acceptable Security Technologies (DeSAST) design tool

The second design tool produced within this project is titled *Designing for Socially Integral Security Technologies* (DeSAST). From Table 6.1 this design tool falls within category 4 - *A malleable methodology where the processes and format of results are both adaptable*. The DeSAST approach seeks to address some of the inherent weakness of the FCC; namely the absence of secondary processes to address the 43 commonalities and the burden this places on the end-user. It operates by producing a different series of questions, activities and outputs tailored specifically to address each of the forty-three identified commonalities. The idea being that by maximising both the *flexibility of process* and the *flexibility of output* and by tailoring each of these specifically to the different commonalities of controversy, DeSAST will be uniquely placed to address the multi-criteria conundrum.

DeSAST is reproduced in its entirety in Appendix *L*. What follows here is a discussion of the tool itself; specifically its design, purposes, and content. This is supplemented in Chapter 6.2.3 by an example *method-for-use*, and lists of anticipated benefits and weaknesses of adopting this design approach.

In relation to the design of DeSAST, perhaps the most obvious feature of the printed version of DeSAST is its large size, being 117 pages long. However, the most important factor here is that 86 of these pages comprise the explicit steps to be undertaken; both the questions to be addressed and the specific information to better comprehend these questions. This focus on practical steps, both by length and depth, is much greater than in the other tools and methodologies examined within my dissertation. The danger of presenting this tool, in its detailed entirety, to a potential STEM end-user is that its size will deter them from using it, especially if they are not already convinced that investing the required effort to complete the tool will add design value. As will be shown in Chapter 6.2.3 one way I have attempted to counteract this is I have designed the tool in such a manner that it can be reduced down to its ostensibly

relevant components depending on the proposed ST[156]. Another possibility is to develop DeSAST into a digital format. This would have the effect of obscuring the entirety of its content behind a constructed user interface, thus potentially negating feelings of being overwhelmed. It would permit easier updates based on the emergence of new STs, would potentially speed up the process of using this tool, and would facilitate the distribution of results.

There also exist a number of purposes for employing DeSAST beyond the initially stated intention of assisting the developers and designers of security technologies in anticipating and mitigating negative societal responses to their technologies upstream in the design process. These purposes are not mutually exclusive, they do not need to be of equal weight to the end-user, nor do they even need to be something intended by the end-user. They include the following:

- A *public relations* exercise by the ST developer. The ST developer does not to *actually* value social acceptability to use DeSAST. They may well prioritise sales and profits, and see DeSAST as a tool for maximising these goals. The irony here is that such a cynical yet rational approach doesn't matter. If the relevant changes to a ST design are made such that it provides security in a socially acceptable fashion, then ultimately the 'true' intentions of the developer are irrelevant.
- A tool for imbuing STEM ST designers with a social conscience. DeSAST moves the end-user beyond the *simple* question they are presented with (and one which they are eminently qualified to answer - i.e. *How can we get X to do Y?*) and forces them to address a question they may not otherwise be as inclined or equipped to do so (i.e. *How can we get X to do Y in a manner which respects fundamental rights, minimises physical and mental harm, respects the rule of law, is financially justifiable, is necessary and proportionate, meets social acceptability standards, is not open to misuse…* etc.).
- DeSAST is a tool for sparking the imagination of designers. You cannot consciously address a problem you did not know existed. By highlighting potential sources of

[156] However, this reduction process does involve certain risks. These are discussed in Chapter 6.2.3 and shown to arise in practice in the application of DeSAST in Chapter 7.2 – specifically in relation to the potential misuse/abuse of the case-study ST.

controversy within a ST design, DeSAST can act as a catalyst for the designer to produce novel solutions to possible problems within a design while maintaining requisite functionality.

Regarding the content of DeSAST (i.e. the information included describing each of the forty-three included commonalities of controversy and their associated questions/tasks), from the perspectives of the social scientist, lawyer, rights-activist, and ethicist, much/some of what is included may appear obvious, self-evident, or even naïve. However, this is to ignore the lack of social science training afforded STEM practitioners and the reality of their work environments when designing and developing STs[157]. It is essential to ensure here that any tool developed meets the five design rules set out in Chapter 6.1; the third of which being that *any produced tool must be usable by the intended end-users*. With this rule in mind DeSAST is entirely designed around the conduct of steps to be undertaken. It may require the end-user seek out data so as to answer each question, however no question presupposes or requires the end-user possess special knowledge or training in social research techniques.

Finally, the content of DeSAST is entirely founded upon broad democratic values; a conscious decision on my part. This affords it the potential to transfer and entrench these values into any ST developed with this design tool. Such a fact could lead to the charge that DeSAST is pushing a particular liberal agenda by seeking to embed (or 'future-proof') these values in *solid* form. Without engaging too deeply in this debate, or seeking to affirm/deny such claims, I would point out that the values promoted within DeSAST (respect for human rights, rejection of discrimination, the provision of security, etc.) are all values *already* embedded within our democratic society. Thus from a values-perspective, it is hard to charge DeSAST with being a radical, destabilising instrument.

[157] Points covered in detail in Chapters 1 & 3

6.2.3 A potential method for using DeSAST

DeSAST begins by having the end-user identify which out of the seven categories of commonalities to include within their assessment of their proposed security technology; these seven categories are reproduced from the FCC. Two options for deciding what to include/exclude are presented to the end-user[158].

The first option is for the end-user to simply include all seven categories. The benefit of this approach is that it will ensure all identified commonalities of controversy are examined, including those whose relevance to the security technology in question only becomes apparent *because* they are examined. Restricting the examination of the security technology to only those categories of commonalities which the end-user *ex ante* considers to be relevant increases the risk of the non-identification of unanticipated problems. The drawback with this approach is that time and resources will most likely be wasted addressing categories which are not relevant.

The second option seeks to narrow the scope of commonalities under investigation by presenting the end-user with a question to answer for six of the seven categories:

1. *Physical or Mental Harm*: Does the security technology possess the potential to cause physical or mental harm to the subject?
2. *Liberties & Human Rights*: Does the security technology possess the potential to infringe upon any of our other human rights when seeking to provide us with security?
3. *Questions of Legality*: Does the potential exist for the legality of this security technology to be questioned, challenged, or merely brought into doubt?
4. *Financial Cost of the ST*: Given the financial cost of the security technology, could questions be raised as to whether or not the security technology represents sound financial investment?
5. *Public & End-User Acceptability*: (no question)[159]
6. *Issues of Functionality*: Could any of the functionality aspects relating to how the security technology actually works (i.e., what it does, what it doesn't do,

[158] See pages 326-327

[159] Category 5: Public & End-User Acceptability is automatically included in all applications of DeSAST

what it's intended to do, and what it is capable of doing) be criticised or called into question?

7. *Safety, Security, Misuse & Abuse*: Could the security technology be misused or abused thereby jeopardising the safety or security of citizens or their property? Additionally would it be possible for attackers to avoid or circumvent your security technology?

To assist the end-user in answering these six questions, each of them is accompanied with between two to five points to consider when making this determination[160]. The benefit of his approach is that it saves time and resources by focussing attention to those categories of commonalities which are most likely to have relevance. The downside is that if the end-user erroneously excludes a category on the basis of this initial assessment, the possibility of not addressing relevant commonalities of controversy increases.

Once the end-user has decided on applicable categories they then complete the relevant sections with the design tool. Guidance on this process is provided within DeSAST[161], with the overarching aim being the production of a list of practical design requirements for incorporation into the future design of the proposed security technology.

Each commonality of controversy falling within those categories being assessed is to be addressed individually. To this end the end-user is presented with information on the controversy followed by a series of tasks to be undertaken. The information component is presented within a short fact sheet describing the precise nature of the identified controversy. Where appropriate, this fact sheet will also include specific challenges facing the end-user when addressing the commonality; these were drawn from the case studies of previous controversial security technologies.

The questions to be addressed differ markedly between commonalities, both by process and output; reflecting the fact that each different series of questions/activities is bespoke to the commonality for which they were created. Ultimately however, both

[160] See pages 324-325
[161] See pages 329-330

the manner in which the end-user chooses to complete each question and the form of the output they produce is entirely at their discretion. They can choose to follow the format/output contained within the questions provided or they can produce their own alternatives if they feel it better fits their project and/or work requirements. Should end-users choose to follow the DeSAST output format, then they will be encouraged throughout the process to identify design requirements to address potential future sources of social controversy.

The anticipated benefits and weaknesses of utilising DeSAST

Anticipated benefits of adopting the DeSAST approach include the following:

a) By incorporating flexible processes and outputs, DeSAST should be well placed to cope with both any security technology and all the identified commonalities of controversy without further modification.
b) The design of the individual questions accompanying each commonality is tailored specifically to that commonality, thereby increasing the relevance of those questions.
c) DeSAST should be usable by all end-users without requiring specific training or qualifications.
d) The information included within DeSAST is designed to highlight social and ethical elements of security technologies which may not otherwise normally receive attention by the anticipated end-users. It both focuses the minds of the developers and educates then for future projects.
e) This tool is designed to be used without public interaction or engagement.
f) DeSAST will produce a list of social design requirements for a security technology before that technology has been produced and avoidable mistakes have become locked in.
g) DeSAST will stimulate imaginative thinking and assist in the formation of creative solutions.

Anticipated weaknesses of adopting DeSAST include the following:

a) Quality of output will depend upon both the effort and foresight ability of those end-users using the tool.
b) The flexibility of processes and outputs may translate into variability of the quality of processes and outputs adopted by the end-users within a DeSAST assessment.
c) The tool is not a silver bullet for guaranteeing socially acceptable security technologies. The fact that no problems are identified through the use of DeSAST does not mean the security technology will not provoke social resistance.

7. Stage 5 – Determining Success for the Created Tools

The activities and findings of the previous four Stages of this research project are as follows:

- ***Stage 1*** involved the completion of case studies of controversial STs which led to the identification of 43 commonalities of controversy.
- ***Stage 2*** involved interviews with STEM practitioners involved in the design of these STs so as to develop an understanding how these individuals operate within their particular field with its unique characteristics.
- ***Stage 3*** combined the output of Stages 1 and 2 to produce assessment criteria. These were applied to off-the-shelf candidates and it was determined that they were unable to meet the aims of this research project.
- This necessitated ***Stage 4*** whereby bespoke design tools were created to assist the developers of *all* STs in addressing *all* the identified commonalities of controversy in their designs so as to identify and mitigate future negative social reactions to their products upstream in the design process.

As discussed in the introduction[162] this structure was not settled upon as the somehow intuitively correct approach before this research project was undertaken; rather the results obtained from each Stage had implications for those Stages which followed, beyond dictating the form of those Stages. For example, had an existing design tool been identified during Stage 3 then it would not have been necessary to produce my own design tools. And as a direct consequence of the time and resources necessarily allocated to creating new design tools in Stage 4, my intentions to validate and improve upon these tools through a detailed iterative process with designers had to be curtailed. This action must now constitute future work in this area.

Despite these circumstances it was still possible to at least begin this validation process by applying one of the created tools (DeSAST) to a ST currently being developed by engineers within University College London. Both the inherent challenges of this validation process and the results obtained are presented below.

[162] See Chapter 1.1.1

7.1 The challenge of validation

Having created two design tools based on the work completed in the earlier stages of this research project, the logical question arising is 'do these tools work?' This ostensibly simple question presupposes the existence of a definitive answer while belying the inherent complexities involved in formulating a valid response. Fundamental challenges to addressing this question include the following:

a) *The subjective nature of measuring success*. The development and implementation of STs involves many different actors, including the manufacturers, governments, end-users, and the public. This last group (the public) is not a homogenous entity; rather it comprises a vast number of actor subgroups as determined by the criteria used to segregate each group[163]. Each of these actors may possess different interpretations of what outcome constitutes success. As an example consider the case of an airport whole-body scanner[164]. A *government* may prioritise security through the creation of a scanner which provides as clear an image of the passenger's body as possible to control what is carried onto an aircraft. The *airport operator end-user* many prioritise speed of operation. The *border control end-user* may prioritise reliability through the absence of false positives/negatives. The *travelling passengers* may individually prioritise different combinations of convenience, speed, security, privacy, adherence to their religious beliefs, equality, etc. When each of these different actors prioritises a different outcome, *success* and *failure* become subjective concepts tied to the beholder. Indeed success for one group may be considered an abject failure by another, and *vice versa*. This situation becomes even more complicated in the absence of any agreed metric for determining success in relation to each of these prioritised outcomes. Even if, for the purposes of this research project, we disregard the

[163] Each subgroup is distinguishable by the combination of criteria its members must possess so as to enjoy membership to that subgroup (e.g. age, sex, ethnicity, medical conditions, etc.), and there is virtually no limit to what each individual criterion is or how they can be combined. As such the total number of subgroups can even outnumber the size of the population as individual citizens can simultaneously belong to multiple subgroups.

[164] See Chapter 2.2 for a detailed case-study.

different subgroups and adopt as our measure of success '*the introduction of a ST into society without social resistance*'[165] other challenges still exist.

b) *The problem of correlation versus causality*. If the developers employ a design tool when constructing a ST, and that ST is subsequently introduced into a society without controversy, without further evidence it is not true to conclude that the ST was acceptable *because* the design tool was employed, as opposed to this acceptance simply being a coincidence. This represents the classic correlation versus causality dilemma which underpins the following scenarios:

    i. *Equating the successful introduction of a ST as a success of the design tool*. Just because a ST created with assistance from a design tool is successfully introduced into society, this does not mean its success is a direct result of this tool. The ST may still have been a success even if the design tool was not used.

    ii. *Equating the unsuccessful introduction of a ST as a failure of the design tool*. Just because a ST created with assistance from a design tool is unsuccessfully introduced into society, this does not mean its failure is a direct result of this tool. The ST may still have been a failure even if the design tool was not used.

As stated in Chapter 7, actions to comprehensively address these issues in relation to the two created design tools so as to validate their efficacy goes beyond the length and time constraints of this research project[166]. Nevertheless, it is both possible and prudent at this point to present a preliminary case-study whereby the *Designing for Socially Acceptable Security Technologies*[167] tool was applied to a ST currently under development; that being a passive Wi-Fi radar device for 'seeing through walls'.

[165] Or at least resistance at a level which does not threaten the continued operation of the ST

[166] However, it is worth noting here that despite the challenges listed in this section there is no need to assume that the design tools will be fixed and static. They can be adapted and improved upon by the end-users in light of problems that arise with their use.

[167] See Appendix *L*

7.2 Single case-study of DeSAST use: through-wall sensing of people by Wi-Fi radar

The technology currently being developed seeks to achieve through-the-wall detection of the movements of people within a building by the passive monitoring of changes in Wi-Fi radio wave frequencies observed when radio waves reflect off moving objects[168]. This technology exploits those Wi-Fi radio signals increasingly present in homes and workspaces; signals constantly being emitted from devices such as wireless internet routers. The current prototype of this technology is about the size of a suitcase, and as it does not emit any radio waves itself it is undetectable (Chetty et al 2012; Clark 2012; Hambling 2012). Reported potential uses for this technology include detecting intruders, the unobtrusive monitoring of children and the elderly, hostage situations, urban warfare, spying, and improving CCTV operation by slaving a directional camera to a Wi-Fi detector. Given these uses this technology readily constitutes a ST.

One of the engineers creating this device agreed to apply the *Designing for Socially Acceptable Security Technologies* (DeSAST) tool to their technology. This application of DeSAST was subject to a number of limitations and qualifications which must be clearly stated.

Firstly no training had been afforded the participant on the use of DeSAST, with the only instructions provided being those contained within the tool itself (see pages 326-330). Secondly due to the engineer's busy schedule they spent less than an hour using the tool; hence they were unable to complete all of those sections and related tasks within DeSAST they had identified as applicable. Thirdly because this interviewee represents a sample size of one, our findings are not generalisable; i.e. I cannot make claims pertaining to the wider population of ST researchers, developers, and designers as I do not know whether our single sample represents a typical or atypical example of this population. Fourthly my proximity to this engineer may have influenced the results. I have known this individual for a number of years and have worked with them before. As a result they may have spent more time on DeSAST and produced more/less candid feedback then would have otherwise been the case.

[168] A phenomenon known as the Doppler effect.

Therefore this single case-study in no way constitutes a comprehensive validation programme. It was not a rigorous application of this design tool, and the results from this single sample cannot be claimed to represent those of the wider population of future STs. And yet, as will be seen from the following discussion, even with these caveats there is considerable value in presenting the responses of this engineer (which constitute the first application of DeSAST within a *live* ST research and design project); if only to provide a starting-point for the future evaluation and development of this and future design tool(s).

7.2.1 Results, discussion, and implications

DeSAST section selection process

After reading the *How To Use The Design Tool* instructions at the beginning of DeSAST[169], for Step 1 the subject chose to apply *option B* whereby they used the guide provided within DeSAST[170] to determine which of the seven sections to include as opposed to examining them all. The resulting selection decisions are presented below in Table 7.1 (the accompanying quotations are the reasoning provided by the engineer when justifying their decisions):

Table 7.1 Selection decisions within trial of DeSAST

1. Physical or mental harm: **rejected** "*There's no harm from radio waves so this doesn't apply*".
2. Liberties and human rights: **applied** "*Privacy is an issue*".
3. Questions of legality: **undecided** "*I have no idea of the legality of it*" (at this point I advised the subject to consider including those sections for which they were uncertain as to relevance)
4. Financial cost of the ST: **rejected** "*It is a low cost technology; that is how we'd sell it. It piggy-backs off existing Wi-Fi and [the Wi-Fi is] the expensive part of it*".
5. Public and end-user acceptability: (always applied)

[169] See pages 326-327
[170] See pages 324-325

6. Issues of functionality: **applied** (no reasoning was provided here by the subject)

7. Safety, security, misuse and abuse: **rejected** "*No, I don't think there's any safety or abuse issues here. The idea [is] it's a covert technology so no-one knows about it. It's not subject to issues of abuse really. And it is safe*".

These selection decisions highlight both the benefits and drawbacks of including an option whereby the user can choose to limit those sections examined when undertaking the tool. While the decision to include *Section 2: Liberties and Human Rights* on the basis of obvious privacy concerns is easily defendable given the nature of the ST, four other selection decisions (for Sections 1, 3, 4, and 7) raise questions about both the wisdom of including an option to omit sections from any use of DeSAST, as well the ability of an end-user to critically examine their own ST.

Section 1: Physical or Mental Harm was quickly rejected for inclusion by the engineer with little deliberation on the basis that radio waves are not physically harmful. However this decision displays a focus on the underlying *science*, rather than consideration of its application within a ST. It is at least conceivable that a sufficiently-sensitive, operational Wi-Fi radar system could produce mental distress (and possibly harm) if it was applied in an oppressive, continuous manner. For example if could be employed by local government authorities to continuously monitor tenants within social housing to ensure they were not sub-letting rooms. Such an intrusion might result in negative mental implications. This example highlights the difficulty in determining how remote any harm arising from the application of a potential ST needs to be before Section 1 is included/excluded from DeSAST.

Section 3: Legality highlights a lack of sufficient guidance within the existing directions to cover the situation whereby the end-user does not know if a section should or should not be included. A conservative approach would dictate the inclusion of those sections whose relevance is uncertain, and the instructions should be amended accordingly.

Section 4: Financial Costs is also interesting in that the decision to reject, which from the provided reasoning is based solely on the initial unit price of their ST, may prove correct despite their apparent failure to engage with the other identified aspects of

this section. For example, what about on-going expenses, do the costs outweigh the benefits, and can the same results be active through cheaper means?

For *Section 6: Functionality* the end-user applied this section without any justifying logic. Given the value derived from examining the justifying statements provided for other sections it may prove useful to require the end-user set down the reason(s) behind their selection decisions. This would certainly assist in any process of results validation by the end-user or their colleagues.

Section 7: Safety, Security, Misuse & Abuse best highlights the danger of rejecting a section because the end-user of DeSAST (i.e. the engineer) fails to independently conceive of problems arising. As will be expanded upon below, Section 7 does raise issues and should have been included. This definitive statement is based on both online comments by individuals concerned about the misuse of this technology[171], as well as recognition by the interviewed engineer that their ST could be abused or misused in a variety of ways after I had challenged them with a number of possible examples.

This outcome is a source of concern. For if the end-user cannot recognise (or cannot be trained to recognise) when their proposed ST may lead to social resistance when applying DeSAST then this tool will not be able to act as a sufficient substitute for external actors[172].

Addressing subsections

By following Step 2 of the instructions[173] the subject examined the sub-sections and associated questions comprising the *Framework for Common Controversies within Security Technologies* (FCC)[174] which forms the underlying structure for DeSAST. They began with *Section 2: Liberties & Human Rights* but quickly found it challenging and somewhat daunting to meaningfully address the questions as presented in this format.

[171] See Table 7.3 below for a selection of comments collected from forums, blogs, and responses to articles about the proposed Wi-Fi technology which appeared online

[172] Based on the interviews conducted with STEM practitioners in Stage 2, suggested reasons as to why engineers close to a technology may have difficulty identifying sources of social resistance include; a lack of social science training within STEM education, and a mind-set focussed on solving technical problems without consideration of the social aspects traditional failure to consider or prioritise social elements within a design. For an expanded discussion see Mitchener-Nissen (2013)

[173] See page 327

[174] See Chapter 6.2.1, with the FCC reproduced in Appendix *K*

After reading through Section 2 their immediate response was "*there's loads of these questions, but the point that's in my head is that [the ST] detects a person but it has no identity information around that person. So I don't know. For me it's hard to say 'do the security benefits outweigh the losses to other rights'*[175]".

At this point I suggested the subject attempt to undertake the detailed steps provided within DeSAST for addressing *Qn.2a. Could the security technology impact someone's right to privacy* of Section 2, given that they had already identified *privacy* as a potential issue for their ST. It was also my hope that the concrete steps comprising this subsection would make the task of addressing the questions posed in the FCC appear more manageable and achievable for the subject. Table 7.2 below is a reproduction of Qn.2a as it appears within DeSAST and the subject's accompanying responses[176].

Table 7.2 Responses to Qn.2a within trial of DeSAST

Qn.2a: Could the security technology impact someone's right to privacy?
Step 1. What new security will the proposed technology provide us with? *Surveillance.* What (if any) privacy will we lose because of the proposed security technology? *In an airport scenario [where we could apply this ST] it makes CCTV more efficient, so you're not really losing anything to be honest. Where CCTV is monitoring it has a narrow field of view. [Our ST] is just monitoring what's going on outside this field of view, and if it detects something it sends the CCTV there; it slaves the CCTV camera. So the privacy is already lost in my view … because the CCTV is a lot more intrusive. So what privacy will we lose? Nothing much.* Show how this technology is justified by balancing this change in security against any change in privacy? *My justification is that CCTV already infringed the privacy rights, and it's got to an accepted level now by people. CCTV is accepted in an airport as it has already been deployed, so all the privacy issues have been agreed and dealt with. So this [my technology] makes CCTV do its job better.*
Step 2. What privacy is lost by people who are not engaged in criminal activities but are subjected to the proposed security technology (i.e. what is the privacy cost to innocent people)? *Okay, so that he same thing [as step1].*

[175] This represents subsection 2f. *Do the security benefits outweigh losses to other rights?* within the FCC, expanded upon in pages 358-359 of DeSAST (see Appendix L)

[176] Minus the text-boxes included within the tool proper – see Appendix *L*, pages 348-349

Step 3. What privacy will citizens be left with if the proposed security technology is successfully introduced? *CCTV invades your privacy so they'll just have a little bit less privacy because CCTV will work better.*
Step 4. What are the possible negative security implications of this privacy intrusion? *The technology could be jammed. You could take a jammer and jam the signal; that's the negative. So if people were relying on it, for it to work to detect motion, and it gets jammed, then they might think it's working and it's not. That could be an issue.*
Step 5. How can you minimise both any impact on privacy and any risk of negative security implications through the design of your security technology? *I think that's a good question. To make sure it does work with CCTV means that it overcomes all those obstacles like privacy. If it's an attachment to an existing technology that makes it [the existing technology] work better – it just feeds information into it to work better – then it should be minimal.*
Step 6. If you were personally being monitored by your proposed security technology, would you be as likely to engage in the following legal activities: having an affair?_____; writing a blog criticising your employer/the government/police?_____; sending/receiving legal documents to/from your lawyer?_____; organising or participating in a protest rally?_____; purchasing, downloading, and/or viewing pornography?_____; seeking online advice on abortions/ medical conditions/assisted suicide?_____; supporting/sending money to Wikileaks?_____. If you answered 'yes' to any of these your technology is having a chilling effect. How can you minimise this effect through your design? (This Step was not addressed by the subject as they did not consider it applicable)

Upon completion of these, the subject was complimentary of both the content in *Qn.2a.* and the overall process of the steps contained therein. Their statements also imply they derived value from undertaking these processes in that now they could self-justify why they believe privacy is not an issue here:

> The privacy issues have already been dealt with because it latches onto an existing technology. So it (DeSAST) has made me think about how I can justify this when discussing privacy, which is good.
>
> I say this is good because in a way now I am thinking about it, and I have gone through these questions and everything, and [to me] it just shows that privacy isn't a huge issue actually.

In my efforts to develop a tool for end-users, and for what I was hoping this tool could achieve, it is this second statement which affords me the greatest encouragement and satisfaction, despite the end-user's flawed conclusions. For I hold that in relation to this research project a design tool justifies its own worth if it possesses the capacity to encourage the developers of STs to at least recognise and contemplate social issues in relation to their technologies where they would not have otherwise done so.

In relation to the responses afforded to each step, there are a number of observations to be made. The first is that the engineer interpreted at least one of the steps (Step 4) in a different manner to that which I had envisaged when creating it. They interpreted that step as focussing on the possible security implications of their ST failing to work. My intention when constructing this step was to focus their attention on the negative implications for an individual's security as a result of their loss in privacy caused by this technology. It is possible that this represents a weakness in the manner I presented/worded that question; although equally it is an arguable a feature of language that all questions are open to multiple interpretations.

Either way, I do not consider this different interpretation by the end-user in any sense constituting failure. Once I have produced this tool it is ultimately for the end-user to interpret and derive value from the steps therein. While I had a particular interpretation for this step in mind when producing it, the alternate interpretation is equally valid, as demonstrated by the results produced. Identifying the threat of jamming as a source of failure leading to decreased security is a logical, legitimate, and valuable outcome for the end-user. It is inescapable that when producing these design tools, just as for technologies, once they are passed on to others to use my control over these tools is diminished or extinguished[177].

Secondly, there will always remain the situation that the manner in which the end-user chooses to address a step will impact their ability to identify possible negative social implications via that and related steps. For example from Step 1 onwards the subject chose to focus on *one particular application* for their ST; that being enhancing CCTVs within airports. For this narrow application they concluded privacy would not be an

[177] By accepting this position I am essentially adopting Latour's *first principle*; that "*the fate of what we say and make [of facts and machines] is in later users' hands*" (Latour 1987, p.29).

issue. But from their final comments they appeared to extrapolate their results to conclude that privacy will not be an issue regardless of the context. This jump from a *narrow* application to *all* applications of their ST is not logically valid. While it is quite possible that the particular airport usage may be considered socially acceptable, it does not follow that other applications will be equally accepted. This point is afforded credence by the negative comments posted about this technology in various online forums, blogs, and in the 'comments' sections to articles describing it. The following is a selection of negative responses:

Table 7.3 Negative comments regarding the proposed Wi-Fi radar ST

Obviously this privacy violation is immoral and will foster criminal activity rather than regulatory since regulation has been defunded. (http://phys.org/)
Great, now are we going to have to "WiFi-Proof" our houses? The military may not be interested in snooping on us, but who knows what lowlife or [private investigator] or other snooper may be keeping track of what's going on in your home. (http://news.cnet.com/)
We have a Constitutional right to privacy. This will be the last straw. (http://www.sodahead.com)
More big brother. (http://www.sodahead.com)
A clear violation of search and seizure. (http://www.sodahead.com)
Good and Bad – if there [is] a guy behind that wall who is about to kill people that is good….If the police are spying on citizens – that is bad. (http://www.sodahead.com)
I don't know about anyone else, but I don't trust the government to use such a device only for 'good.' Also, if the government has it, what's going to stop the 'bad' guys from getting it and also using it? (http://www.sodahead.com)

In spite of these comments, it is readily conceivable that within the specific contexts of airports the use of this Wi-Fi radar to enhance existing CCTV systems may constitute a socially acceptable use for this ST which does not raise specific privacy concerns. If so there is still value within the responses of the subject. What may be useful here is the development of *best practice* guidelines to assist end-users when completing the individual steps within DeSAST to avoid logical traps without stifling creativity.

Thirdly, the iterative process of improving DeSAST is made easier by having end-users apply it in 'real-world' situations. For example, based on the responses of the subject I agree that Step 2 is not sufficiently different from Step 1 to warrant separate inclusion.

Fourthly, by requiring the end-user to physically set-down their responses to each step, these form a valuable source of data for future learning should they underestimate or fail to recognise a source of social resistance within their design which arises after implementation. It also constitutes an information resource which can be used by colleagues to validate or challenge the conclusions of the original end-user.

Considering a rejected section

Given the limited remaining time available to the subject to spend on DeSAST, at this point I directed the discussion away from those sections they had indicated as relevant and onto one of the sections they had rejected as irrelevant; that being *Section 7 Safety, Security, Misuse & Abuse*. This was primarily chosen as a result of comments collected from online discussions whereby individuals indicated concern for the potential abuse of this ST as discussed above.

Given the subject had rejected this section I presented them with hypothetical misuses for this technology so as to challenge their initial decision. Below are the scenarios and subsequent responses:

- A legitimate end-user uses this technology to spy on an ex-partner to check if they are sleeping/living with somebody new: "*There is the husband/wife thing, but what makes it a bit moot for me is that if a husband was to see how many people were in the house, rather than using this technology, for this is not what it's for, it would be a lot easier to put a recorder under the couch or a small camera somewhere as opposed to getting this, and this would be a lot more expensive than those. But I see what you mean*".
- A Council wants to use your technology to determine how many people are living in each of their social housing flats, which may prove a cost effective use of this technology: "*Yeah, good point, yeah. I think that's a really good question actually*".

- A thief wants to use this technology. Could they use it to tell me how many people are inside a bank at the moment, or whether there's a security guard inside a building? "*Yeah, I see what you mean*".

These examples represent one of the fundamental challenges in creating these design tools; that being how to get the end-user to critically evaluate (or 'red-team') their own designs and ideas. As the subject said here; "*I wouldn't have been able to come up with these examples without you [prompting me]*".

This challenge repeated itself when the subject reconsidered Section 7 within the FCC. In the format of a framework they were unable to identify any potential problems for their ST based on the information provided without prompting or challenges by the researcher. This led them to conclude:

> I mean I can see how it [the tool] can stimulate thought to make people think how it can be abused, but like I said, I think I have a lot more insight because you (the researcher) were here. If I was in my office reading this I would just write 'no, no, no' and I think that's [a point worth highlighting].

These results call into question the efficacy of the FCC operating as a design tool. What was not determined here (as it was not applied) is whether the steps provided within Section 7 of DeSAST could act as a suitable substitute for the researcher, and encourage independent critical evaluation by the end-user. If it can do so then determining the optimal combination of design tool and associated training course will be fundamental to success here.

<u>Final observations by the subject</u>

There were other valuable insights provided by the subject during the course of this application of the design tools. The first relates to when a design tool should be employed, and how this changes in the context of who is employing it. Given that the science underpinning Wi-Fi radar is still being developed, the subject stated it was hard to answer questions on the potential capabilities of future STs. This in turn affected their ability to identify possible instances of social resistance. It was their opinion that such tools should be applied by "*an academic in the late stages of development*" (referring to a basic-science research stage) "*or a project manager in industry in the*

very early stages of development" (referring to a stage where this science is then applied to create a piece of engineered technology).

This conceptualisation separates the basic and applied research stages; whereby development is described as moving from the basic science (in this case "*basic physics*") to a "*basic technology to a product*". By doing so the subject is subscribing to a linear model of innovation (see Godin, 2006). While the shortcomings of this linear approach as a depicter of how technological artefacts are produced are well documented [178], this conceptualisation may serve another purpose here. By compartmentalising the *basic* from the *applied* science phases, and by attempting to move the application of DeSAST further downstream to the applied stage, the engineer (justifiably or otherwise) is creating the conditions whereby *they* are not responsible for addressing the tricky ethical and social questions DeSAST is designed to highlight.

Secondly there was an express statement by the subject to the effect that even at this early stage in its development, they could conceive of multiple different ways in which the ST could be designed; "*I have a list of about 10 different ways it can be realised, and each one looks different – has a different architecture, a different set of receivers, different way it looks and operates.*" This represents an incredibly important validation for the approach adopted within this entire research project; that being to focus on the upstream design component of ST development. It confirms that the designer is not operating under a single deterministic vision for their completed ST, thus at least the possibility exists to shape the end-product by mitigating potential sources of social resistance through design requirements.

Finally on DeSAST itself, there was recognition that it could add value to the research and design of STs. However, it is still likely to be challenging for end-users to identify potential sources of future social resistance within this products:

> We are still at the basic science stage [in this Wi-Fi project], and I think this [DeSAST tool] is very applications heavy, but I think it works. It seems to be good in terms of its guidance, but again, I think the misuse stuff (see Section 7 of DeSAST) – I would just say 'how can it be misused?' It's difficult.

[178] See amongst others Kline (1985) and Kline and Rosenberg (1986).

8. Conclusions

This concluding chapter begins by collating and presenting the primary findings of my dissertation. These results comprise my original contributions to the existing empirical research previously undertaken within the field of ST design and development. This is accompanied by discussions both of the limitations of my research and those areas identified as requiring future work.

Following this, and to complete my dissertation, brief discussions on three related topics are included. These being: (1) the dual-use potential of the results produced herein; (2) alternative methods for incorporating *the public* into the design of STs while maintaining secrecy; and (3) a discussion of what my empirical research is able to *say* to the wider field of ST research.

8.1 Statement of results

Results with applicability beyond the borders of my specific dissertation were obtained within each of the diverse stages of the research project. I have collated and summarised the main findings here from each of the individual stages so as to highlight the contribution of my dissertation to the empirical research within the field of ST.

Stage 1: *Case study analysis of controversial STs*

Despite the diversity of form and function existing within the wide range of STs, it is possible to identify common controversies that appear repeatedly across multiple STs. Aided by the recurring nature of these controversies it is also possible to objectively organise them into categories. As a result of this realisation, in Table 2.13 I have been able to produce a taxonomy of common controversies incorporating forty-three controversies with applicability to the design and operation of any ST, irrespective of the nature of the controversy itself.

Stage 2: *Interviews with engineers and scientists*

The primary findings arising from the interviews with STEM practitioners engaged in the design of STs are as follows. Firstly there is the paucity of ethics or social impact

training for the interviewed ST designers. This was highlighted by the absence of ethics and social impact training within their core curriculum, which raises the question of whether concepts such as *engineering ethics* have any practical manifestation at the university level or only exist as an academic ideal? This lack of ethics and social science training is equally noted in those interviewees from disciplines other than engineering such as mathematics, physics, and computer science. Furthermore none of the interviewees undertook electives in social science or ethics. There is also an identified lack of conviction amongst the interviewees that they would opt for a subject on the social impact or ethics of their chosen fields should it be offered. Because there is already so much to learn within their chosen fields, interviewees are predominantly of the opinion that their time would be better spent on their core discipline and/or specialisation.

On the construction and implementation of design tools to assist ST developers in anticipating and mitigating future negative societal responses to their technologies upstream in the design process, the interviewees confirmed that they currently did not use such tools. However, they were in favour of doing so providing they consider these tools add value to the design process.

Stage 3: *Assessment of existing tools/methods for use as an upstream design tool*

A set of eleven assessment criteria for the evaluation of tools/methods as suitable candidates for assisting the upstream development of STs are produced in Chapter 4.1. These represent the culmination of both the case studies undertaken in Stage 1 and the interview process undertaken in Stage 2. They are applied to selection of existing tools/methods in Chapter 5 and remain available for future assessment use.

Stage 4: *Creation of design rules and bespoke tools*

The primary output of Chapter 6 is a set of essential design rules for producing the upstream design tools envisaged within my dissertation[179]. These are essentially the combined product of all the previous chapters and can be used to guide the production of any future upstream design tool for assisting the developers of STs.

[179] These are set out in detail in Chapter 6.1

Also produced within Stage 4 are two design tools produced in accordance with these essential design rules; namely *Designing for Socially Acceptable Security Technologies* and the *Framework for Common Controversies within Security Technologies*. These are reproduced in their entirety in Appendices *K* & *L*.

8.2 Future research work

As discussed in the preface my dissertation is intentionally broad in scope as it is intended to begin the process of addressing largely unexplored areas within the field of ST research. This is especially true in reference to the lack of research aimed at looking across both STs and commonalities of controversy (i.e. the *multi-criteria conundrum* as discussed at the start of Chapter 6.1). As a result there are a number of areas requiring future research work that are highlighted by the results of my dissertation. Additionally, given the dynamic nature of both ST construction and the values of societies, it is necessary to reflexively update many of the results contained herein. As such there would be value in engaging on future work within the areas identified below.

In relation to the *Taxonomy of controversial security technologies*[180], this is not intended to be a static document, but should be added to and refined as new controversial STs emerge. As stated in the case study methodology section (Chapter 2.1.1), my research is cross-sectional, not longitudinal; in that it represents a temporal snap-shot of controversial STs operating at the time the case studies were undertaken. Since this point in time new candidates have arisen with the potential to add to this taxonomy, most notably PRISM and Tempora (see Huhne 2013), while in the future, advances in drones and their use in domestic policing and surveillance operations may well also constitute a viable candidate (see Pilkington 2013).

Additionally, as new assessment methods are developed they should be assessed against the design criteria set out in Chapter 4.1 to determine; (a) whether they can operate as design tools for STs as conceptualised within my dissertation, (b) whether they can be modified to do so, or (c) if neither A or B is possible, whether there are

[180] See Table 2.13

elements of these new methods which can be included in future ST design tools. One promising area which may prove useful in this regard is that of *Responsible Research and Innovation*. Described by von Schomberg (2011, p.9) as:

> a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view on the (ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products (in order to allow a proper embedding of scientific and technological advances in our society),

or more broadly by Stilgoe et al (2013) as "*taking care of the future through collective stewardship of science and innovation in the present*" (p.1570). While obviously relevant to the design of STs, with its emphasis on *inclusion*[181] it remains to be determined how responsible research and innovation can incorporate a requirement for secrecy.

Finally the two tools developed as part of my research (FCC and DeSAST) need to be fully tested. A preliminary application of DeSAST is undertaken in Chapter 7 but this is only intended a precursor to a rigorous assessment and validation process; not a substitute for one.

8.3 The dual-use potential of this dissertation

As stated throughout my dissertation, once a design tool (such as DeSAST or FCC) has been created and made public, the creator of these tools loses the ability to control who uses them and how they are used. Decisions on how they are to be applied now rest in the hands of future users. In an effort to move away from the notion of STEM practitioners utilising the knowledge incorporated within DeSAST and the FCC for the purpose of improving the social acceptability of their STs, I have identified three alternative end-users, each with different agendas.

1. *Campaigners opposed to STs*: Both the FCC and DeSAST could be used by those opposing STs as blueprints for identifying specific shortcomings in the design of a ST. These shortcomings could then be used as focal points for developing social resistance through targeted media campaigns, etc.

[181] The four dimension of responsible innovation proposed by Stilgoe et al (2013) are *anticipation*, *reflexivity*, *inclusion* and *responsiveness*.

2. *Those seeking to circumvent and/or abuse STs*: A number of the commonalities of controversy identified within my dissertation arise from the ability of a ST to be abused/misused or from a ST failing to achieve its purported security goals. It would be possible to use these commonalities to critique an implemented ST and identify weaknesses in its design that would assist in the circumvention of that ST, or to highlight ways to abuse that technology.
3. *Sociologists researching ST designers and developers*: Chapter 7 entailed an initial application of DeSAST by a STEM practitioner to a ST under development. The resulting responses by this end-user were incredibly rich given the small amount of data collected. As a result, I argue that DeSAST could be used by sociologists, when examining how STEM practitioners go about developing STs, as a starting point for the construction and conducting of interviews. It could constitute a ready-made tool for eliciting the thought processes of those designing STs. By identify design elements and potential future controversies which the developer considers important and likely, and those which they reject, could provide valuable insights into the how ST designers operate.

This list of possible end-users and their applications of the FCC and DeSAST is far from comprehensive, and I make no claims over the likelihood these design tools *will* be used in such manners. The point here is simply to argue that human ingenuity affords us the capacity to take an object developed for one purpose and apply it to another; and that my design tools are in no way exempt from this practice.

8.4 Alternative methods for incorporating the public

As stated in Chapter 1.3, in seeking to assist the developers of STs in identifying and mitigating potential sources of social controversy within their designs, I have chosen to work within those existing structures that govern the security industry. As such the dominant factor influencing my work here is that of *secrecy*. As a consequence DeSAST is designed to operate without input from external actors (i.e. the public). However, as I also state in Chapter 1.3, even though I adopt this approach throughout my dissertation I do not endorse processes which prevents the designers of STs benefiting

from the valuable views of the public upstream in ST design. Equally though, per my discussion in Chapter 1.4 of the legitimate need for maintaining elements of secrecy so as to maximise their potential of STs to provide security, neither do I advocate the wholesale abolition of secrecy. The question therefore becomes, if the restrictions on secrecy were to be relaxed (but not abolished) how could input from the public on a future ST be obtained without compromising its potential to *work*? I have two suggestions here. The first is the use of *trusted citizen panels*, and the second is the use of *proxy technologies*.

Trusted citizen panels would consist of bringing together as diverse groups of individuals as possible (essentially a citizen jury) who would examine, and give their opinions of, proposed STs. They would be afforded complete access to the ST with no information being withheld from them. These individuals would volunteer to be members of such panels, though given the obvious security risks such panels would create, all those who are selected would first have to pass some form of security screening. However, so as to ensure these panels are sufficiently removed from the state and the security industry, such that they possess a different standpoint from those who desire the development and introduction of the STs under examination, no member of the panel can be a current/past member of the police, security services, military, etc., or employed as a ST designer. The results of these panel discussions would then be fed back into the ST design process.

The use of proxy technologies for conducting research would work in the following way. Rather than presenting a particular proposed ST for discussion by the public, that ST could be anonymised by breaking it down into a number of elements/capabilities which are then presented as being parts of different hypothetical STs. These could then be assessed by different public focus groups. The anticipated benefit of this approach would be that the public could be involved in the design of STs in a manner which minimises the possibility of secrecy will be compromised. Additionally those members of the public on these focus groups would not need to be security vetted.

Doubtlessly there are potential problems with both of these approaches, but the point of this discussion here is not to engage in a detailed analysis. Rather it is to make the point that with a little lateral thinking there are different ways the views of the public

can be ascertained and infused into the design of STs at only a minimal risk to those technologies. Given the potential damage arising from STs failing to attain public legitimacy, I argue that the benefits of such public engagement outweigh the inherent security costs.

8.5 Points relevant to security technology research

Looking beyond the knowledge generated throughout this project to specifically meet the originally stated research aim, in producing this dissertation my research also raises a number of points with relevance for the wider field of ST research. The first is the need to develop this field from a more holistic perspective. While there is a wealth of past and ongoing research in STs, it is primarily based on silos of knowledge; i.e. specific technologies or classes of technologies (such as CCTV or surveillance technologies), or specific controversies or classes of controversies (such as privacy or human rights). By continuing to focus research into these individual silos (as opposed to a holistic approach in which ST research is not limited by technology or controversy) two consequences arise. The first is the loss of opportunities taking such a narrow focus creates, such as the ability to;

- make connections between different STs,
- identify those factors which are common across STs and subsequently make use of such knowledge, and
- recognise the value that work undertaken in one area may have beyond that area.

The second consequence is that such an approach fails to reflect the development of STs with broader reach. This includes both the creation of *ST systems* (whereby individual STs are linked together) and the creation of STs with previously unachievable reach and functionality (such as the use of one ST as a platform to facilitate the operation of others, or a ST which brings together into a single *black box* STs which in the past would have been applied separately).

The second point arising from my research is the imperative for researchers to gain a greater understanding of how the current STs are developed and to reflect this knowledge when producing methods for engaging with society. This point is primarily

aimed at the variety of innovation and impact assessment models *as they are applied to STs* which fail to take account of the impact of secrecy, and as a result assume a set of social rules/permissions which do not exist. Specifically I am referring to any assessment method which includes engagement with the public (as an identified stakeholder) as one of the steps involved. Failure to do so will result in the production of assessment methods which represent sound theory but cannot be applied to those STs that are designed and operate under constraints of secrecy (which arguably would include the majority of those that would benefit most from such assessment methods).

My final point of relevance pertains to the field of ethics in relation to the design of STs. There is a need to challenge the lack of social training for STEM actors who create these products, beginning with the assumption that their formal education courses incorporate sufficient ethics training. In relation to engineering, there appears to be an element of a 'job well done' sentiment within the academic literature pertaining to the rise of engineering ethics. However, through my interviews with engineers it becomes clear that it is perfectly possible (if not highly likely) that these individuals will complete their formal education without ever addressing ethics or social issues. In addition, STs today are also created by scientists who are even less likely than engineers to encounter these subjects in their formal education; such as physicists, mathematicians, and chemists. While academics in the past have addressed the need for ethics within engineering given its perceived role in the design process of technologies, there is not the same focus on the need to address the same lack of ethics and social science training of mathematicians. Yet in the field of STs, the work of mathematicians may well possess the greatest potential to (negatively) impact the rights of citizens through the development of data mining and profiling algorithms.

8.6 Reflections on the design of socially acceptable STs

My decision to undertake this PhD began with a simple question, '*why is it that we keep building such socially unacceptable STs?*' which in turn led to the reflective retort '*and what exactly can I do about it?*' My response is this dissertation, intended to assist those STEM practitioners who are designing STs to anticipate which of their design

choices are most likely to evoke social resistance so they can choose an alternative path. In so doing I am trying to work with the world as it exists today, accepting the enforced restrictions on public access and participation. These are ideals which are all too often sacrificed in the name of secrecy and national security. I also accept the undeniable fact that when it comes to the overall process of the design and deployment of STs, there are no silver bullets for ensuring public acceptability. From the chain of actors beginning with the state official who desires a particular technology so as to address a perceived security threat through to the end-point at which a technological fix in the form of a ST has been deployed into the hands of multiple end-users, there are simply too many variables and actors at play for a single perfect solution for ensuring acceptability to exist. Incremental measures which seek to address individual links in the chain, such as the design tools aimed at upstream developers that I have proposed, may be the only realistic way forward.

This imperative to provide *security* is often used by government, security services, and law enforcement officials to justify the design and development of some new ST. But there is a point I made in the introduction of my dissertation which needs to be reiterated here; that being the distinction between *legality* and *legitimacy*. A security technology can be conceived, designed, and deployed in a manner which meets all the legal requirements operating within a society. However, if this ST is not considered acceptable by the citizens who comprise this society, then they will not afford it *legitimacy*. As my research has repeatedly shown, if a ST does not attain this status of *legitimacy* then it will not survive unscathed in the mid- to long-term. When this happens everybody involved, including the state, the company producing the technology, the end-user, and those subjected to this unacceptable ST, suffer as a result. This is why the desire to design socially acceptable security technologies is a worthwhile and necessary pursuit.

9. Bibliography

Adams, K. & Jennison, V (2007) What we do not know about police use of Tasers. *Policing*, 30(3), pp.447-65.

Agar, J (2005) Identity cards in Britain: past experience and policy implications. *History & Policy*, Article 33.

Aimeur, E., Brassard, G. & Molins, P (2012) Reconstructing Profiles from Information Disseminated on the Internet. *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on Social Computing (SocialCom)*. DOI: 10.1109/SocialCom-PASSAT.2012.38, pp.875-883.

Aiming for a safer solution (2013) Aiming for a safer solution. *The Job*, January 2013, pp.12-13.

Akrivopoulou, C. & Psygkas, A (eds.) (2011) *Personal Data Privacy and Protection in a Surveillance Era*. United States, IGI Global.

AoL News (2010) *Genital Jokes Lead to Airport Beating*. Last accessed online at http://www.aolnews.com/ca/article/genital-jokes-airport/19469195

Ball, J. & Moore, A (1997) *Essential Physics for Radiographers*. 3rd edition. United Kingdom, Blackwell Science Ltd.

Balmer, B (2012) *Secrecy and Science: A Historical Sociology of Biological and Chemical Warfare*. Farnham, Ashgate Publishing Ltd.

Balton, V. & Stewart, T (2002) *Multiple Criteria Decision Analysis: an integrated approach*. Kluwer Academic.

Barfield, J (2010) *Taser Related Injuries: Fact or Fiction*. Florida, Florida Department of Law Enforcement.

Barrow, C (1997) *Environmental and Social Impact Assessment: An Introduction*. London, Arnold.

Baxter, J (2010) Case Studies in Qualitative Research. In *Qualitative Research Methods in Human Geography (3rd edition)*, edited by Iain Hay. Ontario, Oxford University Press.

BBC News (2006) *Peeping tom CCTV workers jailed*. Last accessed online 15/09/2011 at http://news.bbc.co.uk/1/hi/england/merseyside/4609746.stm

BBC News (2008a) *Spy law 'used in dog fouling war'*. Last accessed online 11/12/2013 at http://news.bbc.co.uk/1/hi/uk/7369543.stm

BBC News (2008b) *Family's shock at council spying*. Last accessed online 05/09/2011 at http://news.bbc.co.uk/go/pr/fr/-/1/hi/england/dorset/7343445.stm

BBC News (2008c) *Council removes mosquito device*. Last accessed online 03/05/2010 at http://news.bbc.co.uk/go/pr/fr/-/1/hi/scotland/tayside_and_central/7787423.stmshed

BBC News (2009a) *'Council 'spying' to be restricted*. Last accessed online 05/09/2011 at http://news.bbc.co.uk/go/pr/fr/-/1/hi/uk_politics/8003123.stm

BBC News (2009b) *Mosquito upsets young travellers*. Last accessed online 30/12/2013 at http://news.bbc.co.uk/go/pr/fr/-/1/hi/england/devon/8419982.stm

BBC News (2010) *Poole council loses school catchment 'spying' tribunal*. Last accessed online 11/12/2013 at http://www.bbc.co.uk/news/uk-england-dorset-10839104

BBC News (2012) *Manchester Airport's Body Scanners Scrapped*. Last accessed online 22/10/2013 at http://www.bbc.co.uk/news/uk-england-manchester-19620981

BBC News (2013) *James McCormick guilty of selling fake bomb detectors*. Last accessed online 16/12/2013 at http://www.bbc.co.uk/news/uk-22266051

Becker, H (1997) *Social Impact Assessment: method and experience in Europe, North America and the Developing World*. London, UCL Press Limited.

Berg, B. & Lune, H (2012) *Qualitative Research Methods for the Social Sciences*. United States, Pearson.

Beynon-Davies, P (2006) Personal identity management in the information polity: The case of the UK national identity card. *Information Polity*, 11 (1), pp.3-19.

Bloomberg (2013) *Naked-Image Scanners to Be Removed From U.S. Airports*. Last accessed online 11/03/2013 at http://www.bloomberg.com/news/2013-01-18/naked-image-scanners-to-be-removed-from-u-s-airports.html

Bok, S (1989) *Secrets: On the Ethics of Concealment and Revelation*. New York, Vintage.

Booth, R (2013) *Fake bomb detector conman jailed for 10 years*. Last accessed online 16/12/2013 from http://www.theguardian.com/uk/2013/may/02/fake-bomb-detector-conman-jailed

Borrion, H., Bouhana, N., Guo, H., Chetty, K., Smith, G., Woodbridge, K. & Baker, C (2008) *Detecting Terrorists: Requirements and system specifications for a radar-based system*. NATO SET-125 Symposium on Sensors and Technology for Defence Against Terrorism. Manheim, Germany.

Bowcott, O (2013) *What are secret courts and what do they mean for UK justice?* Last accessed online 08/11/2013 from http://www.theguardian.com/law/2013/jun/14/what-are-secret-courts

Bradshaw, M. & Stratford, E (2010) Qualitative Research Design and Rigour. In *Qualitative Research Methods in Human Geography (3rd edition)*, edited by Iain Hay. Canada, Oxford University Press.

Braidwood, T (2009) *Restoring Public Confidence: Restricting the Use of Conducted Energy Weapons in British Columbia*. Canada, Braidwood Commission of Inquiry of Conducted Energy Weapon Use.

Brey, P (2011) Anticipatory Technology Ethics for Emerging IT. In *CEPE 2011: Crossing Boundaries* (conference proceedings). Last accessed online 01/08/2013 at http:// users.gw.utwente.nl/Coeckelbergh/site/publicaties/Conference%20Proceedings.pdf

Burghouts, G., den Hollander, R., Schutte, K., Marck, J., Landsmeer, S. & den Breejen, E (2011) Increasing the security at vital infrastructures: automated detection of deviant behaviors. *Proc. SPIE 8019, Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense* X, 80190C (June 02, 2011). DOI:10.1117/12.884579.

Butschi, D., Carius, R., Decker, M., Gram, S., Grunwald, A., Machleidt, P., Steyaert, S. & Van Est, R (2004) The Practice of TA; Science, Interaction, and Communication. In *Bridges between Science, Society and Polity: Technology Assessment - Methods and Impacts*, edited by Michael Decker and Miltos Ladikas. Berlin, Springer.

Callerame, J (2010) *Manufacturer says radiation dose from airport scanner minimal*. Last accessed online 15/06/2010 from http://www.diagnosticimaging.com/news/ display/article/113619/1537426

Casey-Maslen, S (2010) *Non-kinetic-energy weapons termed 'non-lethal'*. Geneva, Geneva academy of international law and human rights.

Cavoukian, A (2009a) *Whole Body Imaging in Airport Scanners: Building in Privacy by Design*. Information and Privacy Commissioner of Ontario.

Cavoukian, A (2009b) *Privacy by Design*. Last accessed online 09/08/2010 at http://www.privacybydesign.ca/content/uploads/2009/01/privacybydesign.pdf

Cavoukian, A (2011a) *The 7 Foundational Principles: Implementation and Mapping of Fair Information Practices*. Last accessed online 09/08/2010 at http://www.ipc.on.ca/images/Resources/pbd-implement-7found-principles.pdf

Cavoukian, A (2011b) *Privacy by Design: The 7 Foundational Principles*. Last accessed online 09/08/2010 at http://www.privacybydesign.ca/content/uploads/2009/08/7 foundationalprinciples.pdf

Cavoukian, A (2013) Privacy by Design: Leadership, Methods, and Results. In *European Data Protection: Coming of Age*, edited by Serge Gutwirth, Ronald Leenes, Paul de Hert and Yves Poullet. Dordrecht, Springer.

Chetty, K., Smith, G. & Woodbridge, K (2012) Through-the-Wall Sensing of Personnel Using Passive Bistatic WiFi Radar at Standoff Distances. *IEEE Transactions on Geoscience and Remote Sensing*, 50(4), pp.1218-1226.

Clark, L (2012) Radar prototype tracks Wi-Fi signals to spy through walls. *Wired*, 06 August 2012. Last accessed online 03/09/13 at http://www.wired.co.uk/news/archive/2012-08/06/wifi-radar

Clifford, N., French, S. & Valentine, G (2010) *Key Methods in Geography*. London, SAGE Publications.

Compound Security (a) (no date). *Mosquito MK4 with Multi-Age*. Last accessed online 03/05/2011 at http://www.compoundsecurity.co.uk/mosquito-mk4-multi-age

Compound Security (b) (no date). *Mosquito Anti-Vandal System (pdf information sheet)*. Last accessed online 03/05/2011 at www.compoundsecurity.co.uk

Corbin, J. & Holt, N (2005) Grounded Theory. In *Research Methods in the Social Sciences*, edited by Bridget Somekh and Cathy Lewin. London, SAGE Publications.

Council for Science and Society (1978) *Harmless Weapons*. United Kingdom, Barry Rose Ltd.

Crawford, A (2009) Criminalizing Sociability through Anti-social Behaviour Legislation: Dispersal Powers, Young People and the Police. *Youth Justice*, 9, pp.5-26.

Cunha, M., de Andrade, N., Lixinski, L. & Féteira, T (eds.) (2013) *New Technologies and Human Rights*. Farnham, Ashgate Publishing Ltd.

Custers, B., Calders, T., Schermer, B. & Zarsky, T (eds.) (2012) *Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases*. Heidelberg, Springer.

Da Veiga, A. & Eloff, J (2007) An Information Security Governance Framework. *Information Systems Management*, 24, pp.361-372.

Daily Mail (2013a) *TSA pulls ALL X-ray body scanners from airports over privacy concerns... but claims they were never a health risk to fliers*. Last accessed online 22/10/2013 at http://www.dailymail.co.uk/news/article-2333685/TSA-removes-ALL-backscatter-X-ray-machines-airports-privacy-concerns.html

Daily Mail (2013b) *Conman sold £13 golf ball finders as bomb detectors: Novelty devices sent to Iraq for £27,000 each*. Last accessed online 26/03/2013 at http://www.dailymail.co.uk/news/article-2289629/Conman-sold-13-golf-ball-inders-bomb-detectors-Novelty-devices-sent-Iraq-27-000-each.html#ixzz2Oe6DP4nc

Davison, N (2007) *The Contemporary Developments of "Non-Lethal" Weapons, Bradford Non-Lethal Weapons Research Project*. Department of Peace Studies, University of Bradford.

Davison, N (2009) *'Non-Lethal' Weapons*. United Kingdom, Palgrave Macmillan.

den Boer, M (2011) Technology-led Policing in the European Union: An Assessment. *Journal of Police Studies*, 2011(3), pp.39-58.

DfT - Department for Transport (2010a) *Assessment of comparative ionising radiation doses from the use of rapiscan secure 1000 x-ray backscatter security scanner*. Last accessed online 30/06/2010 from http://www.dft.gov.uk/pgr/security/aviation/airport /securityscanners/securityscanner/

DfT – Department for Transport (2010b) *Interim Code of Practice for the Acceptable Use of Advanced Imaging Technology (Body Scanners) in an Aviation Security Environment*. Accessed online from http://www.dft.gov.uk/pgr/security/aviation /airport/securityscanners/codeofpractice/pdf/cop.pdf on the 13/05/2011.

DfT – Department for Transport (2010c) *Code of practice for the acceptable use of advanced imaging technology (body scanners) in an aviation security environment – A consultation paper*. Last accessed online 03/08/2010 at http://www.dft.gov.uk /consultations/closed/2010-23/consultation.pdf

DHS - US Department of Homeland Security (2009) *Privacy Impact Assessment Update for TSA Whole Body Imaging*.

Ditton, J (2000) Crime And The City: Public Attitudes towards Open-Street CCTV in Glasgow. *British Journal of Criminology*, 40, pp.692-709.

Doorn, N. & Fahlquist, J (2010) Responsibility in engineering: Towards a new role for engineering ethicists. *Bulletin of Science, Technology & Society*, 30, pp.222-230.

Downs, R (2007) Less lethal weapons: a technologist's perspective. *Policing*, 30(3), pp.358-84.

DPTAC – Disabled Persons Transport Advisory Committee (2010) *DPTAC response – DfT Code of Practice for the acceptable use of advanced imaging technology (body scanners) in an aviation security environment*. Last accessed online 13/05/2011 from http://dptac.independent.gov.uk/pubs/consult/47.htm

Dunn, K (2010) Interviewing. In *Qualitative Research Methods in Human Geography (3rd edition)*, edited by Iain Hay. Canada, Oxford University Press.

ECORYS (2009) *Study on the Competitiveness of the EU security industry*. Rotterdam, ECORYS SCS Group.

Editorial staff (2005) No Such Thing as a Non-Lethal Weapon. *New Scientist*, 185, p.3.

EESC – European Economic and Social Committee (2011) *Opinion of the European Economic and Social Committee on the Communication from the Commission to the European Parliament and the Council on the Use of Security Scanners at EU airports*. COM(2010) 311. Last Accessed online 16/05/2011 from http://www.eesc.europa.eu /?i=portal.en.ten-opinions.15051

EHRC – Equality and Human Rights Commission (2010) *EHRC response - DfT Code of Practice for the acceptable use of advanced imaging technology (body scanners) in an aviation security environment.* Last accessed online 15/05/2011 from

http://www.equalityhumanrights.com/legal-and-policy/consultation-responses/department-for-transport-consultation-code-of-practice-for-the-acceptable-use-of-advanced-imaging-technology-in-an-aviation-security-environment/

Ekblom, P (1999) Can we make crime prevention adaptive by learning from other evolutionary struggles?. *Studies in Crime and Crime Prevention*, 8, pp.27-52.

Elias, B (2010) *Airport and Aviation Security: U.S. policy and strategy in the age of global terrorism*. United States, Auerbach Publications.

European Commission (2010) *Communication from the Commission to the European Parliament and the Council on the Use of Security Scanners at EU airports* COM(2010) 311, Brussels, European Commission.

European Commission (2011) *Aviation security: Commission adopts new rules on the use of security scanners at European airports* IP/11/1343, Brussels 14/11/2011.

European Union (2011) Commission Implementing Regulation (EU) No 1147/2011 of 11 November 2011 amending Regulation (EU) No 185/2010 implementing the common basic standards on civil aviation security as regards the use of security scanners at EU airports.

Farrimond, H (2013) *Doing Ethical Research*. Basingstoke, Palgrave Macmillan.

Fidler, D (2005) The meaning of Moscow: "Non-lethal" weapons and international law in the early 21st century. *International Review of the Red Cross*, 87(859), pp.525-52.

Finn, P. & Horwitz, S (2013) *U.S. charges Snowden with espionage* (21 June 2013). Last accessed online 21/11/2013 at http://www.washingtonpost.com/world/national-security/us-charges-snowden-with-espionage/2013/06/21/507497d8-dab1-11e2-a016-92547bf094cc_story.html

Fish, R. & Geddes, L (2001) Effects of stun guns and tasers. *The Lancet*, 358(September 1), pp.687-88.

Friedman, B (2004) Value Sensitive Design. In *Berkshire Encyclopaedia of Human-Computer Interaction*. Great Barrington, Berkshire Publishing Group.

Friedman, B., Kahn, P. & Boring, A (2008) Value Sensitive Design and Information Systems. In *The Handbook of Information and Computer Ethics*, edited by Kenneth Himma and Herman Tavani. Hoboken, John Wiley & Sons, Inc.

GAO – US Government Accountability Office (2009). *Aviation Security: DHS and TSA Have Researched, Developed, and Begun Deploying Passenger Checkpoint Screening Technologies, but Continue to Face Challenges* (GAO report number GAO-10-128). Washington DC, United States Government Accountability Office.

Gawande, A (2011) *The Checklist Manifesto: How to get things right*. London, Profile Books Ltd.

Geels, F (2007) Feelings of Discontent and the Promise of Middle Range Theory for STS: Examples from Technology Dynamics. *Science Technology & Human Values*, 32(6), pp.627-651.

Gellman, B., Linzer, D. & Leonnig, C (2006) *Surveillance Net Yields Few Suspects* (5 February 2006). Last accessed online 30/12/2013 at http://www.washingtonpost.com /wp-dyn/content/article/2006/02/04/AR2006020401373.html

Gerring, J (2004) What Is a Case Study and What Is It Good for? *The American Political Science Review*, 98(2), pp.341-354.

Gerring, J (2007) *Case Study Research: Principles and Practices*. New York, Cambridge University Press.

Ghosh, T., Prelas, M., Viswanath, D. & Loyalka, S (eds.) (2010) *Science and Technology of Terrorism and Counterterrorism*. 2nd edition. Boca Raton, CRC Press.

Gill, M (editor) (2003) *CCTV*. Leicester, Perpetuity Press Ltd.

Gill, M., Bryan, J. & Allen, J (2007) Public Perceptions of CCTV in Residential Areas: "It Is Not As Good As We Thought It Would Be.". *International Criminal Justice Review*, 17(4), pp.304-24.

Gill, M. & Loveday, K (2003) What Do Offenders Think About CCTV? In *CCTV*, edited by Martin Gill. Leicester, Perpetuity Press Ltd.

Gill, M. & Spriggs, A (2005) *Assessing the Impact of CCTV. Home Office Research Study No. 292*. London, Home Office Development and Statistics Directorate.

Global Security (2011) *Vehicle-Mounted Active Denial System (V-MADS)*. Last accessed online 22/11/2013 from http://www.globalsecurity.org/military/systems/ground/v-mads.htm

Godin, B (2006) The Linear Model of Innovation: The Historical Construction of an Analytical Framework. *Science, Technology, & Human Values*, 31(6), pp.639-667.

Goldstein, J., Angeletti, R., Holzbach, M., Konrad, D. & Snijder, M (2008) *Large-scale Biometrics Deployment in Europe: Identifying Challenges and Threats*. Spain, European Commission Joint Research Centre.

Gower, M (2012) *Biometric Passports*. House of Commons standard note SN/HA/4126 last updated 10 February 2012.

Greenwald, G. & MacAskill, E (2013) *NSA Prism program taps in to user data of Apple, Google and others*. Last accessed online 22/11/2013 at http://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data

Greenwald, G., MacAskill, E. & Poitras, L (2013) *Edward Snowden: the whistleblower behind the NSA surveillance revelations*. Last accessed online 22/11/2013 at

http://www.theguardian.com/world/2013/jun/09/edward-snowden-nsa-whistleblower-surveillance#start-of-comments

Grijpink, J (2006) An assessment model for the use of biometrics. *Computer Law & Security Report*, 22(2), pp.316-319.

Guardian (2009) *Scrap ID cards plan, says David Blunkett*. Last accessed online 25/07/11 at http://www.guardian.co.uk/politics/2009/apr/28/blunkett-id-cards

Guardian (2013) *TSA to remove some body scanners from US airports over privacy concerns*). Last accessed online 11/03/2013 at http://www.guardian.co.uk /world/2013n/jan/19/tsa-revealing-body-scanners-removed

Guest, D (2010) *Western Australian Aborigine Tasered in custody*. Last accessed online 15/08/2011 at http://www.theaustralian.com.au/news/nation/western-australian-aborigine-tasered-in-custody/story-e6frg6nf-1225934071657

Guest, G., Namey, E. & Mitchell, M (2013) *Collecting Qualitative Data: A Field Manual for Applied Research*. United States, Sage Publications.

Gutterman, M (1988) A Formulation of the Value and Means Models of the Fourth Amendment in the age of Technologically Enhanced Surveillance. *Syracuse Law Review*, 39(2), pp.647-736.

Hambling, D (2012) Seeing Through Walls With A Wireless Router. *Popular Science*, August 2012.

Hansard (2010) *House of Commons Debate*. (05/01/2010) 503, part 18.

Hawley, C. & Jones, M (2011a) *Export ban for useless 'bomb detector'*. Last accessed online 30/12/2013 at http://news.bbc.co.uk/1/hi/programmes/newsnight/ 8471187.stm

Hawley, C. & Jones, M (2011b) *UK warns world about useless 'bomb detectors'*. Last accessed online 30/12/2013 at http://news.bbc.co.uk/1/hi/programmes/newsnight /8481774.stm

Hempel, L. & Topfer, E (2009) The Surveillance Consensus: Reviewing the Politics of CCTV in Three European Countries. *European Journal of Criminology*, 6(2), pp.157-77.

Hepple, B (2009) Forensic databases: implications of the cases of S and Marper. *Medicine, Science, and the Law*, 49(2), pp.77-87.

Hinsliff, G (2006) *Brown to let shops share ID card data*. Last accessed online 26/07/11 at http://www.guardian.co.uk/uk/2006/aug/06/idcards.immigrationpolicy

Holstein, J. & Gubrium, J (2004) The active interview. In *Qualitative Research: Theory, Method and Practice (2nd edition)*, edited by David Silverman. London, SAGE Publications.

Home Office (2003) *Identity Cards: A Summary of Findings from the Consultation Exercise on Entitlement Cards and Identity Fraud (CM 6019)*. London, Home Office.

Homeland Security (2006) *Privacy Impact Assessments: Official Guidance*. United States, Department of Homeland Security.

Hudson, R (1999) *The Sociology and Psychology of Terrorism: Who Becomes a Terrorist and Why?*. Washington, Library of Congress.

Huhne, C (2013) *Prism and Tempora: the cabinet was told nothing of the surveillance state's excesses*. Last accessed online 11/01/2014 at http://www.theguardian.com /commentisfree/2013/oct/06/prism-tempora-cabinet-surveillance-state

Human Genetics Commission (2009) *Nothing to hide, nothing to fear? Balancing individual rights and the public interest in the governance and use of the National DNA Database*. United Kingdom, Human Genetics Commission.

Balancing individual rights and the public interest in the governance and use of the National DNA Database

Hunter, P (2005) London terrorist attacks heat up identity card debate and highlights uncertainties over their efficacy. *Computer Fraud & Security*, 7, pp.4–5.

ICM (2006) *No2Id Id Card Survey*. Last accessed online 04/08/2011 at http://www.icmresearch.com/pdfs/2006_july_no2id_id_card_survey.pdf

ICRP – International Commission on Radiological Protection (2007) *The 2007 Recommendations of the International Commission on Radiological Protection*. ICRP Publication 103.

Introna, L. & Wood, D (2004) Picturing Algorithmic Surveillance: The Politics of Facial Recognition Systems. *Surveillance & Society*, 2(2/3), pp.177-198.

Jobling, M. & Gill, P (2004) Encoding Evidence: DNA in Forensic Analysis. *Nature Reviews*, 5, pp.739-751.

Johnson, D. & Wetmore. J (2008) STS and ethics: Implications for engineering ethics. In: *The handbook of science and technology studies (3rd edition)*, edited by Edward Hackett, Olga Amsterdamska, Michael Lynch and Judy Wajcman. Cambridge, Massachusetts, The MIT Press.

Joinson, A., Paine, C., Buchanan, T. & Reips, U (2006) Watching me, watching you: Privacy attitudes and reactions to identity card implementation scenarios in the United Kingdom. *Journal of Information Science*, 32 (4), pp.334-43.

Jonas, J (2011) *Sensemaking on Streams - My G2 Skunk Works Project: Privacy by Design (PbD)*. Accessed at http://jeffjonas.typepad.com/(2011).

Joss, S. & Bellucci, S (2002) Participatory Technology Assessment in Europe: Introducing the EUROPTA Research Project. In *Participatory Technology Assessment:*

European Perspectives, edited by Simon Joss and Sergio Bellucci. Gateshead, Athenaeum Press.

Khan, A (2006) Identity Cards: The Final Nail in the Coffin of Civil Liberties? *Journal of Criminal Law*, 70, pp.139-46.

Kline, S (1985) Innovation Is Not a Linear Process. *Research Management*, 28(4), pp.36-45.

Kline S. & Rosenberg, N (1986) An overview of innovation. In *The Positive Sum Strategy: Harnessing Technology For Economic Growth*, edited by Ralph Landau and Nathan Rosenberg. Washington DC, National Academy Press.

Klitou, D (2008) Backscatter body scanners – A strip search by other means. *Computer Law & Security Report*, 24, pp.316-325.

Kravets, D (2011) Court OKs Airport Body Scanners, Rejects Constitutional Challenge (July 15, 2011). Last accessed online 29/12/2013 at http://www.wired.com/threatlevel/2011/07/court-approves-body-scanners/

Kreimer, S (2004-2005) Watching the Watchers: Surveillance, Transparency, and Political Freedom in the War on Terror. *University of Pennsylvania Journal of Constitutional Law*, 7, pp.133-81.

Kroener, I (2010) *CCTV: a technology under the radar?* Doctoral thesis, University College London.

Lawton, B (2001) *Damage to human hearing by airborne sound of very high frequency or ultrasonic frequency (Health and Safety Executive contract research report 343/2001)*. United Kingdom, Her Majesty's Stationery Office.

Latour, B (1987) *Science In Action*. Cambridge (MA), Harvard University Press.

Latour, B (1996) *Aramis, or, The love of technology*. Cambridge (MA), Harvard University Press.

Lewer, N. & Davison, N (2005) Non-lethal technologies - an overview. *Disarmament Forum*, 2005(1), pp.37-51.

Lewin, C (2005) Elementary Quantitative Methods. In *Research Methods in the Social Sciences*, edited by Bridget Somekh and Cathy Lewin. London, SAGE Publications.

Liberty (2010) *Liberty's response to the Department of Transport's consultation on the Code of Practice for the acceptable use of advanced imaging technology (body scanners) in an aviation security environment*. Last accessed online 17/05/11 from http://www.liberty-human-rights.org.uk/pdfs/policy10/liberty-s-response-to-the-body-scanners-consultation-july-2010.pdf

Lin, Y. & Jones, T (2010) Electronic control devices and use of force outcomes. *Policing*, 33(1), pp.152-78.

London Assembly (2013) *Arming the Met: The deployment of less-lethal weapons in London* (Report by the Police and Crime Committee). London, Greater London Authority.

Longhurst, R (2010) Semi-structured Interviews and Focus Groups. In *Key Methods in Geography (2nd edition)*, edited by Nicholas Clifford, Shaun French and Gill Valentine. London, SAGE Publications.

LSE – London School of Economics and Political Science (2005) *The Identity Project: an assessment of the UK Identity Cards Bill and its implications*. London, London School of Economics and Political Science.

Lyon, D (2006) Airport Screening, Surveillance, and Social Sorting: Canadian Responses to 9/11 in Context. *Canadian Journal of Criminology and Criminal Justice*, 48(3), pp.397-411.

Mackay, D (2003) Multiple Targets: The Reasons to Support Town-centre CCTV Systems. In *CCTV*, edited by Martin Gill. Leicester, Perpetuity Press Ltd.

MacKenzie, D. & Wajcman, J (eds.) (1999) *The social shaping of technology*. 2nd edition. Maidenhead, Open University Press.

Malin, B. & Sweeney, L (2004) How (not) to protect genomic data privacy in a distributed network: using trail re-identification to evaluate and design anonymity protection systems. *Journal of Biomedical Informatics*, 37(3), pp.179-192.

Marks, K (2010) *Police face charges for using Taser 13 times on Aboriginal man*. Last accessed online 15/08/2011 at http://www.independent.co.uk/news/world/Australasia/police-face-charges-for-using-taser-13-times-on-aboriginal-man-2098835.html

Martin, M. & Schinzinger, R (1996) *Ethics in engineering*. New York, McGraw-Hill.

Mathieson, S (2011) *Minister destroys national identity register*. Last accesed online 24/07/11 at http://www.guardian.co.uk/government-computing-network/2011/feb/10/minister-destroys-national-identity-register

McCahill, M. & Norris, C (2003) Estimating the Extent, Sophistication and Legality of CCTV in London. In *CCTV*, edited by Martin Gill. Leicester, Perpetuity Press Ltd.

McKenna, S. & Gong, S (1997) Non-intrusive person authentication for access control by visual tracking and face recognition. In *Audio- and Video-based Biometric Person Authentication: First International Conference, AVBPA'97 Crans-Montana, Switzerland, March 12–14, 1997 Proceedings*, edited by Josef Bigün, Gérard Chollet and Gunilla Borgefors. Berlin, Springer.

Merton, R (1968) *Social theory and social structures*. New York, Macmillan.

Mirror (2010) *Airport worker 'abused scanner'*. Last accessed online 16/05/2011 at http://www.mirror.co.uk/news/latest/2010/03/24/airport-worker-abused-scanner-115875-22135148/

Mishan, E (1976) *Elements of Cost-Benefit Analysis*. 2nd edition. London, George Allen and Unwin Ltd.

Mitcham, C (1997) Engineering design research and social responsibility. In *Technology and values*, edited by Kristin Shrader-Frechette and Laura Westra. Maryland, Rowman & Littlefield Publishers, Inc.

Mitchener-Nissen, T (2010) *Assessing and Improving Whole-Body Scanners through Public Involvement*. Dissertation for Masters of Research in Security Science, submitted September 2010 at University College London.

Mitchener-Nissen, T (2013) Addressing social resistance in emerging security technologies. *Frontiers in Human Neuroscience*, August 2013 (7). DOI: 10.3389/fnhum. 2013.00483

Mitchener-Nissen, T., Bowers, K. & Chetty, K (2012). Public attitudes to airport security: The case of whole body scanners. *Security Journal*, 25(3), pp.229-243.

Morris, J (2007-2008) Big Success or "Big Brother?": Great Britain's National Identification Scheme Before the European Court of Human Rights. *Journal of International and Comparative Law*, 36, pp.443-473.

Mountfield, H. & Gearty, C (2010) *In The Matter Of The Human Rights And Equality Implications Of The Introduction Of Full Body Scanners At Airports*. Report prepared by Matrix Chambers. Last accessed online 13/05/2011 at http://www.equality humanrights.com/.../2010_02_16_body_scanners_in_uk_airports_-_counsels__advice _sml__2_.pdf

Muller, D (2000) Criminal Profiling. *Homicide Studies*, 4 (3), pp.234-64.

National Institute of Justice (1999) *Guide for the Selection of Commercial Explosives Detection Systems for Law Enforcement Applications (NIJ Guide 100-99)*. Washington, U.S. Department of Justice.

NATO - North Atlantic Treaty Organisation (2006) *The Human Effects of Non-Lethal Technologies (RTO Technical Report TR-HFM-073)*. NATO.

Naval Explosive Ordinance Disposal Technology Division (2005) *Test Report: The Detection Capability of the Sniffex Handheld Explosives Detector*. Indian Head, Naval Explosive Ordinance Disposal Technology Division.

NCRP - National Council on Radiation Protection and Measurements (2003) *Screening Of Humans For Security Purposes Using Ionizing Radiation Scanning Systems*, NCRP Commentary No.16. NCRP, Bethesda, Maryland.

New York Times (2013) *Unpopular Full-Body Scanners to Be Removed From Airports*. Last accessed online 11/03/2013 at http://www.nytimes.com/2013/01/19/us/tsa-to-remove-invasive-body-scanners.html?_r=0

Newburn, T (2007) *Criminology*. Devon, Willan Publishing.

Norris, C. & Armstrong, G (1999) *The Maximum Surveillance Society: The Rise of CCTV*. Farnham, Ashgate Publishing Ltd.

NO2ID (no year) *The problems with "ID Cards"*. Last accessed online 15/07/11 at http://www.no2id.net/IDSchemes/whyNot

Okur, M (2008) On Ethical and Legal Aspects of Data Mining. *Journal of Yasar University*, 3(11), pp.1455-61.

Peck, T (2010) *'Teen repellent' buzzing devices may be switched off by law today*. Last accessed online 23/10/2013 at http://www.independent.co.uk/news/uk/home-news/teen-repellent-buzzing-devices-may-be-switched-off-by-law-today-2010048.html

PERF – Police Executive Research Forum (2012) *Critical Issues In Policing Series "How Are Innovations in Technology Transforming Policing?"*. Washington DC, Police Executive Research Forum.

Perrow, C (1984) *Normal Accidents*. New Jersey, Princeton University Press.

Pilkington, E (2013) *US states await key drones decision – and the billions that could follow*. Last accessed online 11/01/2014 at http://www.theguardian.com/world/2013/dec/26/us-states-await-key-drones-decision-billions

Posthumus, S. & von Solms, R (2004) A framework for the governance of information security. *Computers & Security*, 23, pp.638-646.

Press, W (2009) Strong profiling is not mathematically optimal for discovering rare malfeasors. *Proceedings of the National Academy of Sciences of the United States of America*, 106(6), pp.1716-19.

Press, W (2010) To catch a terrorist: can ethnic profiling work? *Significance*, 7(4), pp.164-67.

Privacy International (2004). Mistaken Identity; *Exploring the Relationship Between National Identity Cards & the Prevention of Terrorism (Interim Report)*. London, Privacy International.

Quah, E. & Toh, E (2012) *Cost-Benefit Analysis: Cases and Materials*. New York, Routledge.

Rappert, B (2003) *Non-Lethal Weapons as Legitimizing Forces?*. London, Frank Cass Publishers.

Rappert, B (ed) (2007) *Technology and Security: Governing Threats in the New Millennium*. Houndsmills, Palgrave Macmillan.

Rappert, B. & Balmer, B (2007) Rethinking 'Secrecy' and 'Disclosure': What Science and Technology Studies Can Offer Attempts to Govern WMD Threats. In *Technology and Security: Governing Threats in the New Millennium*, edited by Brian Rappert. Houndsmills, Palgrave Macmillan.

Roberts, D (2011) Technology Is Playing an Expanded Role in Policing. *The Police Chief*, January 2011, pp.72-73.

Rosen, J (2007) The Silver Bullet: Protecting Privacy and Security through Law and Technology. *Proceedings of the American Philosophical Society*, 151(3), pp.291-299.

Rosner, L (2004) *The Technological Fix: How People Use Technology to Create and Solve Problems*. New York, NY, Routledge.

Roy, B (1990) Decision-aid and decision-making. *European Journal of Operational Research*, 45(2-3), pp.324-331.

Roy, B (2005) A Case Against Biometric National Identification Systems (NIDS): "Trading-Off" Privacy Without Getting Security. *Windsor Review of Legal and Social Issues,* 19(45), pp.45-84.

Rubin, H. & Rubin, I (2005) *Qualitative Interviewing: The Art of Hearing Data (2nd edition)*. California, SAGE Publications.

Runnymede (2010) *Ethnic Profiling: The Use of 'Race' in UK Law Enforcement*. London, Runnymede.

Rusconi, E. & Mitchener-Nissen, T (2013) Prospects of functional magnetic resonance imaging as lie detector. *Frontiers in Human Neuroscience*, 7, DOI: 10.3389/fnhum.2013.00594

Schermer, B (2011) The limits of privacy in automated profiling and data mining. *Computer Law and Security Review*, 27, pp.45-52, DOI:10.1016/j.clsr.2010.11.009

Schneier, B. (2006) *Beyond Fear*. United States, Copernicus Books.

Schneier, B. (2008) *Schneier on Security*. Indianapolis, Wiley Publishing.

Schot, J. & Rip, A (1996) The Past and Future of Constructive Technology Assessment. *Technological Forecasting and Social Change*, 54, pp.251-268.

Schreuder, A (2009) *Dutch Debate Use of 'Teen Repellent'*. Last accessed online 05/05/2011 at http://www.spiegel.de/international/europe/0,1518,621025,00.html

Seifert, J (2007) *Data Mining and Homeland Security: An Overview*. Washington, Congressional Research Service.

Sengupta, K (2010) *Head of bomb detector company arrested in fraud investigation*. Last accessed online 23/12/2013 at http://www.independent.co.uk/news/uk/crime/head-of-bomb-detector-company-arrested-in-fraud-investigation-1876388.html

Shapiro, J (2002) *Radiation protection: a guide for scientists, regulators and physicians*. 4th edition. United States, Harvard University Press.

SMH – Sydney Morning Herald (2010) *Airport security screener arrested for retaliating over penis size jokes*. Last accessed online at http:// www.smh.com.au/travel/travel-news/airport-security-screener-arrested--for-retaliating-over-penis-size-jokes-20100507-uhs5.html

Smith, M., Petrocelli, M. & Scheer, C (2007) Excessive force, civil liability, and the Taser in the nation's courts. *Policing*, 30(3), pp.398-422.

Smith, P., Spriggs, A., Argomaniz, J., Allen, J., Jessiman, P., Kara, D., Little, R., Swain, D., Follett, M. & Gill, M (2003) Lessons in Implementing CCTV Schemes: An Early Review. In *CCTV*, edited by Martin Gill. Leicester, Perpetuity Press Ltd.

Solove, D (2008) *Understanding Privacy*. Cambridge (MA), Harvard University Press.

Sprague, O (2007) The Deployment of Taser Weapons to UK Law Enforcement Officials: An Amnesty International Perspective. *Policing*, 1(3), pp.309-15.

Stake, R (2008) Qualitative Case Studies. In *Strategies of Qualitative Inquiry (3rd edition)*, edited by Norman Denzin and Yvonna Lincoln. United States, Sage Publications

Steinbock, D (2005) Data matching, data mining, and due process. *Georgia Law Review*, 40, pp.1–84.

Stilgoe, J., Owen, R. & Macnaghten, P (2013) Developing a framework for responsible innovation. *Research Policy*, 42(9), pp.1568-1580.

Strom, D (2005) *Screening Individuals with Backscatter X-ray Systems* (Last updated February 3rd, 2005). Accessed online from the Health Physics Society.

Sweet, K (2009) *Aviation and Airport Security: terrorism and safety concerns*. 2nd edition. United States, CRC Press.

Sun, X., Wang, H. & Zhang, Y (2012) On the identity anonymization of high-dimensional rating data. *Concurrency and Computation: Practice and Experience*, 24(10), pp.1108-1122.

Taylor, B., Alpert, G. Kubu, B., Woods, D. & Dunham, R (2011), Changes in officer use of force over time: a descriptive analysis of a national survey. *Policing*, 34(2), pp.211-32.

Thai, J (2006) Is Data Mining Ever A Search Under Justice Stevens's Fourth Amendment? *Fordham Law Review*, 74, pp.1731-58.

The Constitution Project (2010) *Principles for Government Data Mining: Preserving Civil Liberties in the Information Age*. Washington, The Constitution Project.

Thiel, G (2000) Automatic CCTV surveillance-towards the VIRTUAL GUARD. *IEEE Aerospace and Electronic Systems Magazine*, 15(7), pp. 3-9.

Thomas III, S., Amara, J., Bjork, R. & Swager, T (2005) Amplifying fluorescent polymer sensors for the explosives taggant 2,3-dimethyl-2,3-dinitrobutane (DMNB). *Chemical Communications*, 41(36) pp.4752-4574.

Thomassen, Ø., Brattebø, G., Heltne, J., Søfteland, E. & Espeland, A (2010) Checklists in the operating room: Help or hurdle? A qualitative study on health workers' experiences. *BMC Health Services Research*, 10, pp.342-374.

Thompson, D (1999) Democratic Secrecy. *Political Science Quarterly*, 114(2), pp.181-193.

Tien, L (2004) Privacy, technology and data mining. *Ohio Northern University Law Review*, 30, pp.389–415.

Travis, A (2008) *MI5 report challenges views on terrorism in Britain*. Last accessed online 30/12/2013 at http://www.theguardian.com/uk/2008/aug/20/uksecurity.terrorism1

TSA - Transportation Security Administration (no date) *Imaging Technology*. Last accessed online 18/06/2010 at http://www.tsa.gov/approach/tech/imaging_technology.shtm

TSA – Transportation Security Administration (2010) *TSA Purchases Additional Advanced Imaging Technology Units (And a Quick Word on Automated Target Recognition*. Last accessed online 23/06/2010 at http://blog.tsa.gov/search/label/backscatter

Tucker, J (2012) The Decision Framework. In *Innovation, Dual Use, and Security: Managing the risks of emerging biological and chemical technologies*, edited by Jonathan Tucker. Cambridge (MA), The MIT Press.

Vilke, G. & Chan, T (2007) Less lethal technology: medical issues. *Policing*, 30(3), pp.341-57.

Vincenti, W (1995) The technical shaping of technology: Real-world constraints and technical logic in Edison's electrical lighting system. *Social Studies of Science*, 25, pp.553-574.

von Schomberg, R (ed) (2011) *Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields*. European Commission, Brussels,

Walsh, C (2008) The Mosquito: A Repellent Response. *Youth Justice*, 8, pp.122-133.

Weinberg, A (1967) *Reflections on Big Science*. Cambridge (MA), MIT Press.

Welch, D (2011) *US death ruling raises concern over Taser use*. Last accessed online 08/08/2011 at http://www.smh.com.au/national/us-death-ruling-raises-concern-over-taser-use-20110728-1i228.html

Whitbeck, C (1998) *Ethics in engineering practice and research*. United Kingdom, Cambridge University Press.

Whitley, E. & Hosein, G (2010) *Global Challenges for Identity Policies*. Basingstoke, Palgrave Macmillan.

WHO - World Health Organisation (2008) *Implementation Manual Surgical Safety Checklist*. 1st edition. Geneva, WHO Press.

Winner, L (1980) Do Artefacts Have Politics? *Daedalus*, 109(1), pp.121-136.

Wired (2013) *Mass. Bill to Ban Data Mining of Student Emails*. Last accessed online 27/03/2013 at http://www.wired.com/insights/2013/02/mass-bill-to-ban-data-mining-of-student-emails/

Wolpe, P., Foster, K. & Langleben, D (2010) Emerging neurotechnologies for lie-detection: promises and perils. *American Journal of Bioethics*, 10, pp.40–48.

Wright, D (2011) A framework for the ethical impact assessment of information technology. *Ethics and Information Technology*, 13, pp.199-226.

Wright, D. & Raab, C (2012) Constructing a surveillance impact assessment. *Computer Law and Security Review*, 28 pp.613-626.

Wright, D. & Wadhwa, K (2012) *A step-by-step guide to privacy impact assessment*. Presentation paper for the second PIAF workshop, Sopot, Poland, 24 April 2012.

Wyatt, S. & Balmer, B (2007) Home on the Range: What and Where is the Middle in Science and Technology Studies? *Science Technology & Human Values*, 32(6), pp.619-626.

Yin, R (2009) *Case Study Research: Design and Methods*. 4th edition. United States, Sage Publications.

Zurawski, N (2010) 'It is all about perceptions': Closed-circuit television, feelings of safety and perceptions of space - What the people say. *Security Journal*, 23(4), pp259-75.

Appendices

Appendix *A* Interviewee consent form.

I agree to be interviewed by the interviewer (Mr Timothy Nissen) as part of his PhD research project into the design of security technologies. The purpose of this interview is; (1) to gain an understanding of how security technologies are produced, including the freedoms and constraints under which you work as part of this process; (2) to understand what information, ethical considerations, and social values you import into the design of security technologies; and (3) to identify the components of your professional education. The overall goal of this research project is the creation of design tools which will ultimately assist you in the identification and mitigation of negative social responses to your security technologies, for use upstream in the design process.

This research project is being undertaken at University College London within the Department for Security and Crime Science. This project is under the supervision of Dr Alex Braithwaite (020 76794986, alex.braithwaite@ucl.ac.uk) and Dr Brian Balmer (020 76793924, b.balmer@ucl.ac.uk). If you have any queries or complaints about the conduct of this research project you can contact either of these individuals.

As the interviewee you have the right to refuse to answer any questions during the interviews. You also have the right to withdraw from this project at any time. If you do choose to withdraw after your interview, all data collected from you will be deleted from the dissertation and destroyed.

After preliminary discussions with a number of potential interviewees regarding privacy concerns, the interviewer agrees to the following provisions in relation to this project:

- I will strictly maintain the anonymity of all interviewees, both during and after the completion of this project. This will include:
    - Not including or divulging the interviewee names, their employers, or current/past projects undertaken by the interviewees to anyone – including my supervisors, examiners, or UCL.
- Within the dissertation write-up, strict anonymity will be maintain, including the removal of all identifiers, no use of pseudo-anonymisation, and no use of identifiers that will permit the aggregation of responses.
- All data collected will be stored as encrypted files on a non-networked computer.
- After the interview data is collected and written-up, all interviewees will be afforded the opportunity to vet what is written so as to agree any changes they may require.
- After the write-up is agreed, all collected data will be permanently destroyed.

Signed (the interviewer) ……..

Timothy Nissen

Appendix *B* Extract from Identity Cards Act 2006

Identity Cards Act 2006

1 The National Identity Register

(3) The statutory purposes are to facilitate, by the maintenance of a secure and reliable record of registrable facts about individuals in the United Kingdom-

(a) the provision of a convenient method for such individuals to prove registrable facts about themselves to others who reasonably require proof; and

(b) the provision of a secure and reliable method for registrable facts about such individuals to be ascertained or verified wherever that is necessary in the public interest.

(4) For the purposes of this Act something is necessary in the public interest if, and only if, it is –

(a) in the interests of national security;

(b) for the purposes of the prevention or detection of crime;

(c) for the purposes of the enforcement of immigration controls;

(d) for the purposes of the enforcement of prohibitions on unauthorised working or employment; or

(e) for the purpose of securing the efficient & effective provision of public services

(5) In this Act "registrable fact", in relation to an individual, means-

(i) information about occasions on which information recorded about him in the Register has been provided to any person

Appendix *C* Extract from Identity Documents Act 2010

Identity Documents Act 2010

2 Cancellation of ID cards etc

(1) No ID cards are to be issued by the Secretary of State at any time on or after the day on which this Act is passed.

(2) All ID cards that are valid immediately before that day are to be treated as cancelled by the Secretary of State at the end of the period of one month beginning with that day.

3 Destruction of information recorded in National Identity Register

The Secretary of State must ensure that all the information recorded in the National Identity Register is destroyed before the end of the period of two months beginning with the day on which this Act is passed.

Appendix *D* Marketing material for the ADE651

Original image removed for copyright reasons from this electronic version.

Appendix *E* The origin and nature of the individually identified controversies from each of the twelve controversial security technologies

Code	Origin Of Controversy				Nature of Controversy								Key words, phrases, concepts
	Policy Decisions		End Users	Design Features	C1 Health	C2 Legality	C3 The Public	C4 Rights & Liberties	C5 Cost	C6 Safety & Security	C7 Functionality	C8 Use & Misuse	
	Rules Governing Use	Wider Policy Decisions											
Whole Body Scanners													
WBS1				X		X							illegality
WBS2				X		X		X					Human Rights Act; privacy; proportionality
WBS3	X					X						X	rule of law; arbitrary use
WBS3a		X				X							rule of law
WBS3b		X					X						unnecessary; unjustified
WBS3c	X					X	X						undeterminable illegality; trust
WBS3d	X						X					X	trust; lack of information
WBS3e		X					X					X	trust; lack of oversight
WBS3f	X					X		X				X	discrimination; trust; propensity for abuse; minority burden
WBS4	X					X		X					discrimination; minority burden
WBS5	X			X		X		X			X		discrimination; minority burden
WBS6				X	X						X		unjustified risk; unproven; questionable necessity
WBS6a				X	X								unknown health effects
WBS6b				X	X								unknown health effects
WBS7			X				X					X	trust; misuse
WBS8			X	X			X				X	X	trust; unnecessary functionality; misuse
WBS9	X	X					X	X					trust; unnecessary; unjustified; privacy
WBS10		X					X						trust; unjustified
WBS11	X						X						trust; unfair; minority burden

WBS11a		X				X	X					X	undeterminable illegality; trust; propensity for misuse
WBS12				X				X			X		privacy; excessive functionality
National Identity Schemes													
NIS1		X					X		X				trust; oversold; dubious claims; questionable accounting
NIS2		X		X				X			X		function creep; excessive functionality; privacy
NIS2a		X				X	X				X		data misuse; trust; function creep
NIS3		X		X			X	X					data misuse; privacy
NIS3a	X			X				X				X	privacy; informational self-determination; propensity for misuse
NIS3b	X		X	X		X	X	X				X	data misuse; undeterminable illegality; informational self-determination; privacy; propensity for misuse
NIS3c	X		X									X	propensity for abuse
NIS4				X							X		excessive functionality; questionable necessity
NIS4a				X				X			X		privacy; excessive functionality
NIS4b				X			X	X				X	privacy; freedom to protest; trust; chilling effect; potential for misuse
NIS4c				X				X					privacy; freedom of actions; freedom of expression; freedom of thought; freedom of association
NIS5		X						X					privacy
NIS5a		X					X	X					unjustified; disproportional; privacy
NIS5b	X		X	X		X	X	X		X		X	data misuse; trust; privacy; informational self-determination; jeopardises individual safety; lack of information; potential for misuse
NIS6		X		X			X			X		X	trust; jeopardises security; potential for misuse
NIS7	X							X				X	discrimination; potential for abuse
NIS8				X						X	X		easy to avoid; ineffective
NIS8a				X						X	X		easy to circumvent; ineffective
NIS9				X					X		X		cost; unproven technology; excessive errors; ineffective
NIS9a				X							X		excessive errors; potentially ineffective; decreasing reliability
NIS10		X					X						oversold; inflated public opinion

NIS11		X					X						falling public opinion
NIS11a	X	X		X			X						unacceptability to society
NIS11b		X					X	X					unacceptability to society; unjustified; unfair; unnecessary; privacy
NIS11c				X		X	X	X		X		X	data misuse; trust; privacy; jeopardises safety/security; propensity for misuse/abuse
NIC11d		X		X					X				cost; questionable accounting
NIS11e		X					X		X				unnecessary to society; falling public opinion costs outweigh benefits
NIS12				X						X			jeopardises safety/security
NIS12a			X	X						X		X	jeopardises safety/security; propensity for abuse/misuse
NIS13				X			X			X		X	minority burden; risk rests with citizens; jeopardises safety/security; propensity for abuse/misuse
NIS14		X					X						burdensome to citizens
NIS14a				X			X						burdensome to citizens; risk rests with citizens
NIS15		X							X				cost
National Identity Registers													
NIR1				X									
NIR1a				X					X				costly
NIR1b				X							X		burden outweighs benefits; disproportional effect; questionable necessity
NIR1c		X									X		burden outweighs benefits; questionable necessity
NIR2	X		X			X	X	X		X		X	data misuse; trust; public opinion; privacy; jeopardises safety/security; propensity for abuse/misuse
NIR3		X		X			X	X			X		trust; disproportional; unnecessary; unacceptable to society; privacy; informational self-determination; unnecessary functionality; excessive functionality; burdensome to citizens
NIR3a		X		X			X			X		X	privacy; jeopardises safety/security; propensity for abuse/misuse
NIR3b		X		X			X			X		X	privacy; jeopardises safety/security; propensity for abuse/misuse
NIR4		X		X		X	X			X	X	X	data misuse; privacy; jeopardises safety/security; risk to citizens; doesn't work; propensity for abuse/misuse
NIR5		X	X			X				X		X	data misuse; open to insiders; propensity for abuse/misuse
NIR6	X		X			X	X	X		X		X	data misuse; privacy;

													informational self-determination; trust; risk to citizens; jeopardises safety/security; propensity for abuse/misuse
NIR7	X	X	X			X	X	X		X		X	data misuse; trust; risk on citizens; privacy; informational self-determination; jeopardises safety/security; open to insiders; propensity for abuse/misuse
NIR7a		X		X		X	X	X		X		X	data misuse; trust; risk on citizens; privacy; informational self-determination; jeopardises safety/security; propensity for abuse/misuse
NIR8				X						X			jeopardises safety/security
NIR9			X	X						X		X	jeopardises safety/security; open to insiders; propensity for abuse/misuse
NIR10		X	X							X		X	jeopardises safety/security; open to insiders; propensity for abuse/misuse
NIR11		X				X		X					undetermined legality; privacy
NIR11a	X					X		X					undetermined legality; privacy
NIR11b	X					X		X					undetermined legality; privacy
NIR11b1	X										X		ineffective
NIR11b2	X										X		ineffective
NIR11b3	X										X		unnecessary
NIR11c	X					X		X					data misuse; privacy; informational self-determination
NIR12	X					X							data misuse
NIR12a	X					X		X			X		data misuse; informational self-determination; function creep
NIR12b	X					X		X			X		data misuse; informational self-determination; excessive functionality
NIR12c	X					X					X		data misuse; excessive errors; burden on citizens
NIR12d	X					X		X			X		data misuse; privacy; informational self-determination; excessive functionality
NIR12e	X					X		X		X	X	X	data misuse; privacy; jeopardises safety/security; lacking essential functionality; propensity for abuse/misuse
NIR13	X					X							potential illegality; rule of law
National Identity Cards													
NIC1			X	X						X		X	jeopardises safety/security; open to insiders; facilitates

													crimes; propensity for abuse/misuse
NIC1a				X						X			jeopardises safety/security
NIC1b				X								X	propensity for misuse
NIC1c		X					X			X			oversold; dubious claims; jeopardises safety/security
NIC1d		X								X			jeopardises safety/security; facilitates crimes
NIC2				X						X			ease of circumvention
NIC2a				X							X		doesn't work (doesn't fulfil purpose)
NIC3				X						X	X		jeopardises safety/security; unnecessary functionality (risk causing design feature)
NIC4				X						X			jeopardises safety/security; easy to circumvent; open to insiders
NIC5		X											(na)
NIC5a				X									jeopardises safety/security (creates a target); open to insiders
NIC5b				X							X		burdensome to citizens; (open to) excessive errors; decreasing reliability
NIC6		X									X		questionable necessity
NIC6a		X									X		doesn't work; not suitably effective
NIC6b		X									X		doesn't work (doesn't fulfil purpose); not suitably effective
NIC6c		X									X		doesn't work (doesn't fulfil purpose); not suitably effective
NIC7			X					X				X	discrimination; propensity for abuse/misuse
NIC8			X							X		X	jeopardises safety/security; propensity for abuse/misuse
NIC9		X					X	X	X		X		trust; cost; privacy; excessive functionality
National/Mass Biometric Systems													
MB1				X							X		variable effectiveness
MB1a				X							X		variable effectiveness
MB1b				X							X		variable effectiveness; quality control
MB1c				X							X		quality control
MB1d				X							X		variable effectiveness
MB1e				X							X		variable effectiveness
MB2				X				X			X		discrimination; variable effectiveness; doesn't work (excludes certain groups); minority burden
MB3				X				X			X		discrimination; variable effectiveness; doesn't work (excludes certain groups); minority burden

MB4				X							X		variable effectiveness; not suitably effective; burdensome to citizens
MB5							X	X			X		minority burden; unfair; discrimination; burdensome to citizens; doesn't work (excludes certain groups)
MB6				X						X	X		easy to circumvent; ineffective
MB7				X							X		false positives/negatives
MB8				X			X				X		trust (end-users); oversold (end-users); unproven; doesn't work; burdensome to citizens
MB9				X						X	X		can be circumvented; not suitably effective (limited in what it can achieve)
MB10				X							X		burdensome to citizens
Profiling Technologies													
PT1				X							X		doesn't work
PT2				X							X		ineffective
PT2a				X							X		ineffective (easily circumvented)
PT3				X							X		doesn't work; ineffective (limited effectiveness)
PT4				X		X		X			X		data misuse; discrimination; minority burden
PT5				X		X		X					discrimination
PT5a				X			X						unacceptable to society
PT5b				X		X		X					rule of law; discrimination/inequality
PT5c				X			X	X					minority burden; discrimination (inequality)
PT5d				X			X						unacceptable to society; unfair; minority burden
PT5e				X			X						minority burden; counterproductive
PT6				X				X					discrimination; (facilitates) inequality
PT7				X				X					discrimination; (exacerbates) inequality
PT8		X					X						unfair; minority burden
PT8a		X					X						unfair; unjustified
PT8b		X					X	X					unfair; minority burden; discrimination
PT8c		X					X						unfair; unacceptable to society
PT9				X			X				X		unfair; unjustified; false positives
PT10		X		X			X	X					counterproductive; unacceptable to society; burden outweighs benefits; discrimination
PT11		X								X	X		jeopardies security; burden outweighs benefit;

													counterproductive
PT12		X					X	X					minority burden; discrimination/inequality
PT13				X				X			X		discrimination; false positives; burdensome to citizens; not suitably effective; excessive errors
PT13a				X							X		not suitably effective; burdensome to citizens
PT13b				X							X		counterproductive; burden outweighs benefits; ineffective
Data Mining													
DMI1		X		X		X	X	X				X	data misuse; unacceptable to society; disproportionate; trust; privacy; informational self-determination; jeopardises safety/security; propensity for abuse/misuse
DMI1a		X					X						unacceptable to society
DMI2			X	X			X						trust (end-user); false positives
DMI3		X						X					privacy; liberty; burden outweighs benefits
DMI4	X	X					X						unjustified; unfair; minority (individual) burden
DMI5		X				X	X						undeterminable illegality; unfair
DMI5a		X				X	X						undeterminable illegality; trust (opaque decision making)
DMI5b		X				X	X					X	undeterminable illegality; trust; unfair; minority (individual) burden; lack of information
DMI6		X										X	lack of oversight
DMI6a		X		X								X	lack of oversight; lack of information
DMI6b		X					X					X	trust; lack of oversight
DMI6c		X				X	X					X	undeterminable illegality; (possible) discrimination; trust; unacceptable to society; lack of information
DMI7				X		X							undeterminable illegality; rule of law (unchallengeable decision-making)
DMI7a				X		X							undeterminable illegality; rule of law (unchallengeable decision-making)
DMI7b				X		X						X	undeterminable illegality; lack of information; propensity for abuse/misuse
DMI8		X				X							undeterminable illegality
DMI9				X							X		not suitably effective (correlation not causality)

DMI10		X				X						X	rule of law (fair hearing); indeterminable illegality; lack of information
DMI11				X			X				X		(potentially) unfair; not suitably effective (questionable decision making)
DMI11a				X							X		not suitably effective (beyond limits of technology)
DMI12				X							X		not suitably effective (beyond limits of technology)
DMI13				X							X		questionable accuracy; susceptible to errors; decreasing reliability
DMI13a				X							X		questionable accuracy; questionable data
DMI13b				X							X		questionable accuracy; questionable data
DMI13c				X							X		questionable accuracy; questionable data
DMI13d				X							X		questionable accuracy; questionable data
DMI13e				X							X		questionable accuracy; human error
DMI13f				X							X		questionable accuracy; questionable data
DMI14				X							X		doesn't (always) work (dependent upon uncontrollable factors)
DMI14a				X							X		doesn't (always) work (dependent upon uncontrollable factors)
DMI15		X		X				X			X		liberty; questionable accuracy; not suitably effective (beyond limits of technology)
DMI16			X			X		X				X	rule of law; presumption of innocence; misuse
DMI17				X			X				X		minority (individual) burden; unfair; false positives
DMI17a		X					X				X		minority (individual) burden; unfair; false positives; excessive functionality
DMI18				X							X		questionable accuracy; questionable data; minority burden
DMI18a				X							X		human (programming) error/bias
DMI19		X						X					(na)
DMI19a				X				X					chilling effect; freedom of expression; freedom to protest
DMI19b		X						X					privacy; limiting autonomy; freedom to dissent; chilling effect
DMI19c		X	X			X		X					data misuse; privacy;

													informational self-determination
DMI19d				X				X					privacy
DMI19e		X	X					X		X			privacy; jeopardises safety/security
DMI19f		X						X					privacy; burdens outweigh benefits; chilling effect
Data Matching													
DMA1				X							X		false positives; increasing errors; questionable accuracy/data
DMA1a	X			X			X				X		unfair; disproportional; increasing errors; burdensome to citizens
DMA1b	X	X					X						unfair; minority (individual) burden; disproportional
DMA1c	X	X					X						unfair; minority (individual) burden; disproportional
DMA1c1	X	X					X						unfair; minority (individual) burden; disproportional
DMA2		X		X			X				X		unacceptable to society; disproportionate; trust; dubious claims; questionable accuracy/data; doesn't work
DMA3	X					X	X						undeterminable legality; rule of law; trust
DMA3a	X	X				X	X				X		undeterminable legality; trust; lacking essential functionality
DMA4		X		X		X		X					rule of law (fair trial); undeterminable legality; right to a fair hearing
DMA5		X						X					chilling effect; freedom to protest; freedom of expression/thought/association
Closed Circuit Television													
CTV1		X					X						oversold; doesn't fulfil claims; falling public opinion
CTV2				X			X				X		doesn't fulfil claims; ineffective
CTV2a		X					X				X		doesn't fulfil claims; ineffective
CTV3				X							X		(sometimes) ineffective
CTV3a				X							X		(sometimes) doesn't work
CTV3b				X		X							(sometimes) doesn't meet legal requirements
CTV4		X					X				X		oversold; ineffective (on its own)
CTV5		X					X						oversold; doesn't fulfil claims
CTV6		X					X						(some uses) unacceptable to society; conditional public support
CTV7	X	X					X					X	unacceptable to society;

													conditional public support; disproportional
CTV8			X							X		X	jeopardises safety/security; propensity for misuse/abuse; misuse
CTV9		X				X							illegality; data misuse
CTV10		X					X						unacceptable to society; disproportional
CTV10a		X					X					X	unacceptable to society; propensity for misuse/abuse
CTV10b		X					X						unacceptable to society
CTV11		X	X				X						trust; unacceptable to society; minority burden
CTV11a		X	X				X						trust; unacceptable to society; minority burden
CTV12		X	X				X	X					trust; unacceptable to society; minority burden; privacy/liberty; discrimination/inequality
CTV12a			X				X	X					trust; unacceptable to society; minority burden; privacy/liberty; discrimination/inequality; counterproductive
CTV12b			X				X						trust; unacceptable to society; minority burden; counterproductive
CTV12c			X				X						trust; unacceptable to society
Hand-Held Explosive Detectors													
HED1				X							X		doesn't work (scam)
HED1a				X			X			X			trust; jeopardises safety/security
HED1b				X						X			jeopardises safety/security
HED2				X			X			X			trust; jeopardises safety/security
Mosquitos													
MS1			X				X						unacceptable to society
MS2		X				X							(potential) illegality
MS3		X				X							(potential) illegality
MS4		X				X							(potential) illegality
MS5		X						X					potential human rights violations
MS5a		X						X					freedom of association; freedom of assembly
MS5b		X						X					discrimination (age)
MS6		X				X							illegal
MS6a		X				X		X					illegal; human rights violation
MS6b		X					X						minority burden; counterproductive
MS6c				X	X								(potential) health hazard
MS6d		X					X				X		minority burden; not suitably effective
MS7		X		X		X		X					illegality; discrimination (age);

													freedom of movement; freedom of assembly
MS8		X					X						counterproductive; unfair
MS9		X		X			X				X		disproportional; unfair; minority burden; not suitably effective
MS10		X					X						counterproductive; unfair
MS11				X	X			X					physical discomfort; degrading; discrimination/inequality
MS12		X						X					freedom of movement; discrimination/inequality
MS12a		X				X		X					undetermined legality; discrimination; freedom of movement; freedom of assembly
MS12b		X					X						unjustified
MS13		X						X		X			freedom of assembly/movement; jeopardises safety/security
MS13a		X								X			jeopardises safety/security
MS14		X		X		X		X					undetermined illegality; discrimination (age)
MS14a		X					X						minority burden
MS15		X					X						disproportional; unacceptable to society; minority burden
MS15a		X						X					discrimination/inequality; morally bereft
MS15b		X					X				X		unfair; unjustified; not suitably effective; lacking essential functionality
MS16				X							X		not suitably effective; indiscriminate
MS16a				X			X				X		unfair; indiscriminate
MS17				X			X				X		unfair; discriminatory; indiscriminate
MS18				X			X				X		unfair; not suitably effective
MS18a				X			X						unacceptable to society; unjustified
MS18b				X			X						unfair; unacceptable to society; minority burden
MS18c	X			X			X						unacceptable to society; unjustified; minority burden
MS19		X					X						counterproductive; unacceptable to society
MS19a		X				X	X						banned; counterproductive; unacceptable to society
MS20				X	X								unjustified health effect; pain
Less Lethal Weapons													
LW1		X					X						conditional public support
LW1a		X					X						(acceptable = public support)
LW1b		X					X						unacceptable to society; unjustified
LW2		X					X				X		unacceptable to society; function creep

LW3			X				X						unacceptable to society; unjustified (overuse)
LW4		X			X		X	X					pain; unacceptable to society; unethical
LW5		X	X				X						unacceptable to society; trust; falling public support
LW5a		X					X	X					unacceptable to society; unjustified; counterproductive; freedom to protest
LW6	X						X						unacceptable to society (overuse); counterproductive
LW7				X	X								can kill
LW7a		X					X						trust
LW7b		X					X						trust; falling public support
LW8	X			X	X								causes permanent injuries
LW9		X			X		X						undetermined health risks; unacceptable to society; trust
LW9a		X					X						trust
LW10				X	X		X				X		undetermined health risks; trust; quality control problems
LW11		X					X						trust
LW12				X	X								variable/uncontrollable health effects; minority burden
LW12a				X	X								variable/uncontrollable health effects; minority burden
LW12b				X	X								variable/uncontrollable health effects; minority burden
LW12c				X	X								variable/uncontrollable health effects; minority burden
LW12d				X	X								variable/uncontrollable health effects; minority burden
LW12e				X	X								variable/uncontrollable health effects; minority burden
LW13	X						X						unacceptable to society (targets)
LW14		X			X								risk of injury/death (testing phase)
LW15				X							X		indiscriminate; counterproductive
LW15a			X	X							X		counterproductive
LW16			X			X	X			X		X	undetermined illegality; unacceptable to society; jeopardises safety; misuse/abuse (by officials)
LW17			X				X					X	unacceptable to society; misuse/abuse (by officials)
LW18				X		X	X	X				X	illegality; unacceptable to society; torture; misuse/abuse
LW19	X									X		X	jeopardises safety; lack of oversight; propensity for abuse/misuse
LW20	X		X				X					X	unacceptable to society; conditional public support;

													misuse/abuse (overuse)
LW21	X		X				X	X		X		X	unacceptable to society; minority burden; counterproductive; discrimination; jeopardises safety/security; misuse/abuse (by officials)
LW22	X						X						unacceptable to society; unjustified; counterproductive
LW22a	X		X				X					X	unacceptable to society (use); undermining public support; misuse/abuse (overuse)
LW23	X		X				X					X	counterproductive; misuse/abuse (overuse)
LW23a	X						X						unacceptable to society; conditional public support
LW23b		X					X						(na)
LW24	X			X			X				X		unacceptable to society; disproportional; not suitably effective; counterproductive
LW25		X					X						conditional public support
LW25a		X					X						(dependent on) public opinion; conditional public support
LW26	X											X	propensity for abuse/misuse (overuse); lack of oversight; lack of information
LW26a	X											X	propensity for abuse/misuse (overuse); lack of oversight
LW27	X						X						conditional public support (rules of use)

Appendix *F* The Stage 2 questionnaire

Education questions:

A1: What academic and professional qualifications do you have?

- are you currently studying for any others?
- (if self-taught) how did you go about doing that?

A2: What professional work experience do you have?

(depending on number/type of qualification, repeat question A3 for each)

A3: Describe to me the make-up of your undergraduate / postgraduate course?

A4: Give me an idea of the subjects that you were taught?

- Was it rigid and structured or was there flexibility in the choice of subjects you could take?
- What electives did you choose?
- Were there any compulsory or elective subjects that could be described as social-science subjects, such as the role of the engineer in society, the responsibilities of engineers, the impact of engineering on society, that sort of thing?
    - What were they?
    - (if elective) Did you take them? (why? / why not?)
    - (if taken) What sorts of things were you taught?
- Did topics such as these form smaller components of some of the technically-focussed subjects within your course?
    - What were they?

A5: What about engineering ethics; was this covered/taught?

- (if no social-science / ethics components were taught) In your opinion why is it that issues like 'the role of engineers in society' or 'engineering ethics' were not taught during your degree?
- A7(if social-science / ethics components were offered but not chosen by this interviewee) Why did you choose not to take these subjects? What was your thinking on the matter?

A6: Are you a member of any professional bodies related to your qualifications?

- (if yes) Are you bound by certain codes of conduct as a result of your membership?
  - Do you know what they are?
  - How were you taught/informed about them?

Working Practice questions:

B1: What's your occupation?

- What does this entail?

B2: Talk we through a normal work day for you, assuming one such exists?

B3: Explain to me what happens when a new design project is announced?

- Are you part of the negotiations or tendering process for such design projects or do you only find out about it once others within your company have agreed to do the work?
- What do the discussions leading up the start of accepting a project entail?
- Do you work alone when working on a project or as part of a team?
- How is the work for these projects divided?
  - Is all the work given to one person, or is it divided amongst a number of individuals?
  - Or is all the work given to one team and that team then divides the work?
  - Or is it divided up between multiple teams with each responsible for their own component?
  - (if not just one individual)
    - How does communication occur between the different people / groups working on a project?
    - Is it formal with specified meetings, informal in its occurrence, or a combination of the two?
- Are you given information about the whole of the project you are working on, or just the part of the project that you are specifically working on?
- So from the moment a project has been announced and the work divided, how is the work completed and overseen as the project progresses?
- What happens when or if a previously unrecognised or unanticipated problem becomes apparent when completing a project?

B4: How detailed are the specs you are given at the start of a project?

- How binding are these?
  - Do they allow for 'wriggle-room' or are the parameters always stringently defined and adhered to?
- What happens if you realise there is a better way of doing something which the client may not have realised but will require the specs be changed?
- How much influence do you have on projects you are given?
- Do specs include 'desirable components' as well as 'mandatory components' when working to time and cost constraints?

B5: When working to a budget within a project, how do you justify and rank the costs by importance?

- Is this information decided when the project is negotiated, or does it fluctuate as the project progresses?

Thoughts and Opinions of the Tools:

C1: What tools do you already use to identify social issues when beginning a project?

(for each listed) Explain how they work, how long they take, describe the output, how useful are they, and what are the strengths and weaknesses?

C2: Do you think there needs to be different tools specifically for group-use and others specifically for use by individuals?

- (if yes) Why? And how should they differ?

C3: When would be the best time in the design process to use the proposed tools?

- Why?

C4: How long should it take to learn how to use a tool for the first time?

C5: Regarding the instructions for the tools:

- How long could they be before you stopped reading them?
- Should they be very detailed or should they be fairly abstract?
- Would it be valuable if they contain concrete examples of how each aspect of the tools are applied to a real-life or hypothetical design project?
- Do you think there is a risk that providing specific examples of use would not just provide guidance to the user but would restrict the user by narrowing their perception of how the tool could be used?
- At the point when reading the instructions for the first time, what would make you throw them away and give up on the whole idea entirely?

C6: Would you want to have somebody step you through how the tools are used, especially if you had never used them before?

C7: How long should any tool take to use?

- Is it realistic to think this time would be available?

C8: What format would you want the tool to be in? (paper based or digital)

C9: Which of the following options for the form of this tool appeal to you most?

- Firstly for tools aimed at individuals, would you want them to be based on:
  - Checklists
  - Risk assessment tools, such as event tree / fault trees, etc.
  - Frameworks
  - Games
  - Stories based on real life events
  - Probabilities
  - Some other basis (please state)
- Are there any other designs of tools for individuals you think would make a suitable template?
- Secondly for tools aimed at groups, would you want them to be based on:
  - Checklists
  - Risk assessment tools, such as event tree / fault trees, etc.
  - Frameworks
  - Games
  - Stories based on real life events
  - Probabilities
  - Some other basis (please state)
- Are there any other designs of tools for groups you think would make a suitable template?

C10: Would they need to be bespoke to a particular security product (such as CCTV) or could their design be generalised so as to apply to all products?

- Would it be better to have a generalised base with add-on components depending on the specific technology(s) being looked at (e.g. imaging/recognition, biometric, profiling, etc.)?

C11: Would they need to be bespoke to a particular engineering discipline (such as electronic, civil, or software) or could their design be generalised so as to apply to all disciplines?

- Would it be better to have a generalised base with add-on components depending on the specific discipline(s) of the user?

C12: What do you think this tool should be able to achieve?

C13: What form should the output take so as to be of greatest value:

- Should it be quantitative (a number or value) allowing the identification of potential issues, or should it be qualitative (offering advice or identifying future problems)?
    - Should it be both?
- Should it offer potential solutions based on previous controversies from other similar or related technologies?
    - (if yes) Do you worry this might have the unintended consequence of stifling innovation by focussing designers on the past rather than looking for new solutions?

C14: Regardless of what form the output takes, what would you want a tool like this to actually tell you about what it is you are working on?

- What is would it have to achieve / produce to make you want to use it?
- Would you use these tools if they were mandatory within you project?
- Would you use these tools if they were optional within your project?

C15: What would stop you or prevent you from using the tools I am describing?

C16: Do you see any value in using such tools? (why? / why not?)

C17: Do you think your final products would be improved by the use of such tools?

- (why? / why not?)

C18: Do you think engineers in general would benefit from such tools?

- Do you think they need such tools?

C19: Do you see it as the responsibility of people in your profession to have to worry about any future negative social impacts of designs you are employed to build? ➢ why? / why not? o (if not) primarily whose responsibility do you think this is then?
Final Miscellaneous Questions:
D1: Is there anything else I should be aware of, or keep in mind, when designing these tools?
D2: Is there anything else you wish to tell me?

Appendix *G* Ethical impact assessment of information technology framework[182]

Ethical Principles	Related Values / Issues
Respect for autonomy (right to liberty)	• Dignity • Informed consent
Nonmaleficence (avoiding harm)	• Safety • Social solidarity, inclusion and exclusion • Isolation and substitution of human contact • Discrimination and social sorting
Beneficence	• Universal service • Accessibility • Value sensitive design • Sustainability
Justice	• Equality and fairness (social justice)
Privacy and data protection	• Collection limitation (data minimisation) and retention • Data quality • Purpose specification • Use limitation • Transparency (openness) • Individual participation and access to data • Anonymity • Privacy of personal communications: monitoring and location tracking • Privacy of the person • Privacy of personal behaviour

[182] Taken from Wright (2011)

Appendix *H* Anticipatory technology ethics checklist[183]

- **Harms and risks**
  - Health and bodily harm
  - Pain and suffering
  - Psychological harm
  - Harm to human capabilities
  - Harms to society

- **Rights**
  - Freedom
    - Freedom of movement
    - Freedom of speech and expression
    - Freedom of assembly
  - Autonomy
    - Ability to think one's own thoughts and form one's own opinions
    - Ability to make one's own choices
    - Responsibility and accountability
    - Informed consent
  - Human dignity
  - Privacy
    - Information privacy
    - Bodily privacy
    - Relational privacy
  - Property
    - Right to property
    - Intellectual property rights
  - Other basic human rights as specified in human rights declarations (e.g., to life, to have a fair trial, to vote, to receive an education, to pursue happiness, to seek asylum, to engage in peaceful protest, to practice ones religion, to work for anyone, to have a family, etc.)
  - Animal rights and animal welfare

- **Justice (distributive)**
  - Just distribution of primary goods, capabilities, risks and hazards
  - Non-discrimination and equal treatment relative to age, gender, sexual orientation, social class, race, ethnicity, religion, disability, etc.
  - North-south justice
  - Intergenerational justice

[183] Taken from Brey (2011)

- Social inclusion

– **Well-being and the common good**
  - Supportive of happiness, health, knowledge, wisdom, virtue friendship, trust, achievement, desire-fulfilment, and transcendent meaning
  - Supportive of vital social institutions and structures
  - Supportive of democracy and democratic institutions
  - Supportive of culture and cultural diversity

Appendix *I* Dual-use decision framework[184]

Original image removed for copyright reasons from this electronic version.

[184] Taken from Tucker (2012), p.69

Appendix *J* Assessments of existing methodologies

Candidate: CHECKLISTS (5.2.1)			**Category:** N/A
Employment of other methods: N/A			
Criterion	**Can the candidate meet this criterion?**	**Could it be made to do so?**	**Comments**
1. For use before building commences	Yes		Checklist-centric design tools could readily be created for use before work on the actual construction of a ST commences.
2. Results intended for senior use	Yes		No obvious preclusion here.
3. Cannot involve external actors	Yes		There is requirement of involving external actors when using checklists.
4. Produces design specifications, not policies	Yes		Checklists could assist designers by clarifying the elements of a proposed future design, thus highlighting areas of concern.
5. Output must add value	Yes, but limited in scope	Not obviously	The binary nature of the output from a checklist does possess the potential to add value to any ST. However because of the closed-ended nature of this output the level of detail it can provide to end-users is limited. This in turn could impact upon a checklist design tool's ability to influence the design process.
6. Usable in any workplace	Yes		No obvious preclusion here.
7. No prerequisite expertise required	Yes		Checklists are probably the most intuitive to use of all the methodologies and models examined here.
8. Minimal use-time required	Yes		Because of their simple nature, and the minimal data-input required when using one, checklists should be able to minimise the use-time of a design tool.
9. Addresses multiple, diverse controversies	No	Not without adding secondary processes	The closed-ended nature of a checklist means it will struggle to deal with controversies which are not themselves presented in a 'yes'/'no' format. And given the contextual nature of the vast majority of controversies identified, it is likely a model or methodology possessing the potential to collect more granular information will be required to supplement a checklist-based approach.
10. Adaptable to all	Yes, but	Yes, but	A separate bespoke checklist will need to be created

STs	requiring work	requiring work	for each commonality of controversy. Additionally others may need to be created for each ST those commonalities are applied to. This may result in the necessary creation of an exponential number of checklists.
11. Not based exclusively on 'yes'/'no' answers	No	Yes, by adding secondary processes	Checklists are essentially closed-ended questions; hence the value of any output may be limited. It would be possible to expand upon the nature of this output but this would demand the addition of secondary processes into a checklist-based model which would come into effect depending upon the answer(s) provided to the initial 'yes'/'no' question.

Notes: Checklists have numerous benefits allowing them to successfully meet many of the identified assessment criteria. They can be relative quick, are intuitive to use, do not require external actors, and can readily be employed at the start of an ST project. However there are drawbacks. The nature of their output is limited, which in turn may limit the value they can add to a project. They will struggle to address commonalities which require open-ended responses, and a single defined checklist may not be applicable to all STs without modification.

Conclusion: Checklists could readily be employed within design tools aimed at the developers of STs as envisioned within the constraints of this project, however not as a stand-alone tool. They would either need to supplement, or be supplemented by, other methodologies or models to overcome their intrinsic limitations.

Candidate: PRIVACY & SURVEILLANCE IMPACT ASSESSMENTS (5.2.2)			**Category:** IMPACT ASSESSMENTS
Employment of other methods: Utilises *Stakeholder Engagement*			
Criterion	**Can the candidate meet this criterion?**	**Could it be made to do so?**	**Comments**
1. For use before building commences	Yes		Providing there exists a sufficiently clear vision of the intended ST (i.e., how and where it is to operate, its functionality, etc.) then these impact assessments could be used at the start of a design project.
2. Results intended for senior use	Yes		No obvious preclusions here
3. Cannot involve external actors	No	Possibly	Stakeholder engagement (which includes the public affected by a ST) is an intrinsic component of privacy and surveillance impact assessments. While it would be possible to carry out these assessments without this engagement; (i) the value of the output may be severely undermined, and (ii) it would require the creation of alternative methods of data collection to

			replace the now excluded stakeholder engagement.
4. Produces design specifications, not policies	Yes		These assessment methods identify both risks and possible solutions, thus providing usable output for designers.
5. Output must add value	Yes		By identifying risks and possible solutions for redress during the design stage, the output adds value.
6. Usable in any workplace	Yes		No obvious preclusions here.
7. No prerequisite expertise required	Not necessarily	Perhaps with sufficient resource allocation	Depending on the size of an organisation, and the expertise of its staff, the ability to carry out a sufficiently rigorous privacy/surveillance impact assessment may require either specialised training for staff members or the employment of specialists.
8. Minimal use-time required	No	Not without impacting quality	A full privacy/surveillance impact assessment as it currently exists can take considerable time to carry out. It is not obvious that STEM practitioners working within a ST design environment would be able (or willing) to devote sufficient attention to these models.
9. Addresses multiple, diverse controversies	No	Not obviously or efficiently	Surveillance and privacy assessments are primarily focussed on a small subset of the 43 identified commonalities of controversy. They could not obviously be modified to address commonalities with no reference to surveillance and/or privacy without completely modifying the nature of these assessments – and even if they were this would not be an efficient or recommended use of resources.
10. Adaptable to all STs	'Yes' to those with privacy or surveillance related elements	Not obviously or efficiently	As for Criterion 9 above, they could not obviously be modified to address STs with no reference to surveillance and/or privacy without completely modifying the nature of these assessments – and even if they were this would not be an efficient or recommended use of resources.
11. Not based exclusively on 'yes'/'no' answers	Yes		No obvious preclusions here.

Notes: Privacy and surveillance impact assessments rely on an environment of openness and the inclusion of stakeholders. They can also require considerable resources to conduct, and depend on the possession of specialised skills/training by those conducting them.

Conclusions: These models are not appropriate for use as design tools within the constraints of this research project. Furthermore modifying these models so as to meet the assessment criteria would require fundamentally changing the nature of these models to the extent that they would probably no longer be recognisable as examples of the genre.

Candidate: ETHICAL IMPACT ASSESSMENT FRAMEWORK FOR INFORMATION TECHNOLOGIES (5.2.2)			**Category:** IMPACT ASSESSMENTS
Employment of other methods: A hybrid of *Frameworks* and *Stakeholder Engagement*			
Criterion	**Can the candidate meet this criterion?**	**Could it be made to do so?**	**Comments**
1. For use before building commences	Yes		No obvious preclusions here.
2. Results intended for senior use	Yes		No obvious preclusions here
3. Cannot involve external actors	No	Yes	As devised by Wright, one of the three steps involved in this ethical assessment requires stakeholder engagement to generate debate[185]. This step could be removed from the process, with the negative effects mitigated by expanding upon the framework as has currently been presented[186].
4. Produces design specifications, not policies	Yes		The output is designed to identify mitigating measures for implementation prior to deployment of the ST.
5. Output must add value	Yes		By mitigating identified issues before deployment value is added to the final ST.
6. Usable in any workplace	Depending on resources	Yes	The stakeholder-engagement step as currently devised would require access to both citizens and existing networks of experts. This may be beyond the resources of individual designers or very small teams. However removing this engagement requirement, as per Assessment Criterion 3 above, would probably ensure the tool was usable within any workplace.
7. No prerequisite expertise required	No	Yes	Expertise in engagement methods outlined in Step 2 earlier would be required. However this would be negated by the removal of this engagement process.
8. Minimal use-time required	Possibly	Yes	Removal of the engagement processes would help ensure the use-time requirements of this ethical assessment process were kept to a manageable level.
9. Addresses multiple, diverse	No	Yes	The current model has been developed for information technologies, hence including requirements such as malfeasance[187] which would not

[185] See Chapter 5.2.2.
[186] See Appendix *E*
[187] See Appendix *E*

controversies			translate to controversies arising from less-lethal weapons. However, this could be addressed by modifying the framework produced without the need to begin again.
10. Adaptable to all STs	No	Yes	As per Criterion 9 above.
11. Not based exclusively on 'yes'/'no' answers	Yes		No obvious preclusions here.

Notes: The current ethical impact assessment framework leading into a series of questions to be addressed, employs stakeholder engagement, and is designed around information technologies. However by removing this engagement requirement and expanding the scope of the framework and the questions asked, all of the Assessment Criteria above could be catered for.

Conclusions: The approach adopted here of using a framework coupled with questions to be addressed could be used as the basis for design tools capable of dealing with all of the commonalities identified within this project as well as applicable to any ST.

Candidate: ANTICIPATORY TECHNOLOGY ETHICS (5.2.2)			**Category:** IMPACT ASSESSMENTS
Employment of other methods: Incorporates *Checklists*, *Technology Assessments*, *Forecasting*, and *Expert Engagement*			
Criterion	**Can the candidate meet this criterion?**	**Could it be made to do so?**	**Comments**
1. For use before building commences	Yes		Results from anticipatory technology ethics (ATE) can be employed to guide how a technology is developed.
2. Results intended for senior use	Yes		No obvious preclusions here.
3. Cannot involve external actors	No	Yes, but with an impact on the range of views included	The first stage of ATE (the forecasting stage[188]) involves collecting input from a wide range of actors including engineers, historians, sociologists, marketing experts, etc. It would be possible to modify this stage to only involve STEM practitioners from within a ST design team. However this would obviously impact upon the types of views expressed given the radical reduction in the scope of voices heard.

[188] See the anticipatory technology ethics subsection of Chapter 5.2.2.

4. Produces design specifications, not policies	Yes		ATE possesses the power to produce information to guide how a technology is developed, as well as recommendations for policy makers. The later however, does not preclude the former.
5. Output must add value	Yes		Guidance on the development of a technology as well as the identification of ethical issues all possesses the ability to add value to a final ST.
6. Usable in any workplace	No	Yes	The requirements under the current forecasting stage of ATE may make this tool unusable within all workplaces. However by modifying this requirement, as per Criterion 3 above, it would be possible to maximise the usability of this assessment tool.
7. No prerequisite expertise required	No	Yes	As per Criterion 3 above, modifying the forecasting stage would largely remove the need for prerequisite expertise.
8. Minimal use-time required	Somewhat	Yes	ATE requires end-users undertake a number of steps which together become quite time consuming. However the changes mooted above to the forecasting stage would assist in bringing this down to a more manageable level.
9. Addresses multiple, diverse controversies	No	Yes, by widening the ethical checklist	By modifying and expanding upon the ethical checklist used for cross-referencing those elements of the technology identified by the engineers, it would be possible to account for all of the 43 commonalities of controversy identified.
10. Adaptable to all STs	No	Yes	This should not be a problem once the ethical checklist component is expanded.
11. Not based exclusively on 'yes'/'no' answers	Yes		No obvious preclusions here.

Notes: The wide range of actors included within the presented ATE model contravenes the fundamental requirement of not including external actors. Without modification ATE would not be an appropriate method for meeting the requirements of this research project. However there is value to be taken from the approaches adopted within ATE.

Conclusions: There are significant elements of ATE which could be employed within future ST design tools; namely the role of engineers within the process, and the use of ethical checklists.

Candidate: BASIC FRAMEWORKS (5.2.3)			**Category:** FRAMEWORKS
Employment of other methods: N/A			
Criterion	**Can the candidate**	**Could it be**	**Comments**

	meet this criterion?	made to do so?	
1. For use before building commences	Yes		Can be designed to meet this requirement.
2. Results intended for senior use	Yes		Can be designed to meet this requirement.
3. Cannot involve external actors	Yes		Can be designed to meet this requirement.
4. Produces design specifications, not policies	Yes		Can be designed to meet this requirement.
5. Output must add value	No	Yes	A framework provides the skeleton onto which information and tasks need to be added so as to transform the framework into a practical tool. This is entirely possible, and its effectiveness here will depend on how the framework is shaped and padded.
6. Usable in any workplace	Yes		Can be designed to meet this requirement.
7. No prerequisite expertise required	Yes		Can be designed to meet this requirement.
8. Minimal use-time required	Yes		Can be designed to meet this requirement.
9. Addresses multiple, diverse controversies	Yes		Can be designed to meet this requirement.
10. Adaptable to all STs	Yes		Can be designed to meet this requirement.
11. Not based exclusively on 'yes'/'no' answers	Yes		Can be designed to meet this requirement.

Notes: Frameworks need to be built upon if they are to be transformed into a practical tool. As such all of the Assessment Criteria here can be met. However, it would require the producer of a framework-based tool deliberately choosing to build their tool in a manner compliant with these requirements. Otherwise this completed framework-centric tool would not be appropriate.

Conclusions: A framework is an incredibly useful model, and can form the underlying structure of potentially powerful design tools for meeting the requirements of ST design. They are however only the underlying structure; hence they need to be built upon.

Candidate: EXPANDED FRAMEWORKS (5.2.3)		**Category:** FRAMEWORKS	
Employment of other methods: Virtually unlimited for *expanded frameworks*			
Criterion	**Can the candidate meet this criterion?**	**Could it be made to do so?**	**Comments**
1. For use before building commences	Dependent upon additional processes	Yes, providing additional processes do not preclude this	Can be designed to meet this requirement. Success here depends entirely upon the nature and construction of the additional processes added on to the basic framework.
2. Results intended for senior use	Dependent upon additional processes	Yes, providing additional processes do not preclude this	Can be designed to meet this requirement. Success here depends entirely upon the nature and construction of the additional processes added on to the basic framework.
3. Cannot involve external actors	Dependent upon additional processes	Yes, providing additional processes do not preclude this	Can be designed to meet this requirement. Success here depends entirely upon the nature and construction of the additional processes added on to the basic framework.
4. Produces design specifications, not policies	Dependent upon additional processes	Yes, providing additional processes do not preclude this	Can be designed to meet this requirement. Success here depends entirely upon the nature and construction of the additional processes added on to the basic framework.
5. Output must add value	Dependent upon additional processes	Yes, providing additional processes do not preclude this	Can be designed to meet this requirement. Success here depends entirely upon the nature and construction of the additional processes added on to the basic framework.
6. Usable in any workplace	Dependent upon additional processes	Yes, providing additional processes do not preclude	Can be designed to meet this requirement. Success here depends entirely upon the nature and construction of the additional processes added on to the basic framework.

		this	
7. No prerequisite expertise required	Dependent upon additional processes	Yes, providing additional processes do not preclude this	Can be designed to meet this requirement. Success here depends entirely upon the nature and construction of the additional processes added on to the basic framework.
8. Minimal use-time required	Dependent upon additional processes	Yes, providing additional processes do not preclude this	Can be designed to meet this requirement. Success here depends entirely upon the nature and construction of the additional processes added on to the basic framework.
9. Addresses multiple, diverse controversies	Dependent upon additional processes	Yes, providing additional processes do not preclude this	Can be designed to meet this requirement. Success here depends entirely upon the nature and construction of the additional processes added on to the basic framework.
10. Adaptable to all STs	Dependent upon additional processes	Yes, providing additional processes do not preclude this	Can be designed to meet this requirement. Success here depends entirely upon the nature and construction of the additional processes added on to the basic framework.
11. Not based exclusively on 'yes'/'no' answers	Dependent upon additional processes	Yes, providing additional processes do not preclude this	Can be designed to meet this requirement. Success here depends entirely upon the nature and construction of the additional processes added on to the basic framework.
Notes: Expanded Frameworks refer to frameworks which have been built upon, thus transforming them into practical tools. As a result while all of the Assessment Criteria here can be met, this will depend upon the producer of a framework-based tool deliberately choosing suitable additional process to add onto the bare frame that is the Basic Framework. **Conclusions:** Expanded Frameworks can form design tools for assisting the developers of STs in a manner compliant with those restrictions identified as existing within the security field. However, this is entirely dependent upon the nature of the attached additional processes.			

Candidate: VALUE SENSITIVE DESIGN (5.2.4)	**Category:** DESIGN-FOCUSSED APPROACHES

Employment of other methods: Incorporates *stakeholder engagement*, and elements of *cost-benefit analysis*			
Criterion	**Can the candidate meet this criterion?**	**Could it be made to do so?**	**Comments**
1. For use before building commences	Yes		Providing sufficiently detailed information on the proposed ST exists at the start of the project.
2. Results intended for senior use	Yes		No obvious preclusions here
3. Cannot involve external actors	No	Possibly	This model relies upon identifying the value-impacts of the proposed technology on identifiable stakeholders. Currently this requires stakeholder engagement. If this process cannot be substituted for an alternative which elicits the same information without engaging with these external actors then it will not be able to meet this assessment criterion.
4. Produces design specifications, not policies	Yes		A strength of this model is that it makes explicit the trade-offs of different possible designs.
5. Output must add value	Yes		No obvious preclusion here based on Criterion 4.
6. Usable in any workplace	Possibly	Probably	If the external engagement requirement is removed then this criterion should be achievable.
7. No prerequisite expertise required	Probably yes		Given the nature of the steps involved, it is probable that designers could undertake then but they may require more extensive training than for some of the other models discussed here.
8. Minimal use-time required	Yes		Longer than some other models, but still within acceptable limits, especially once the external involvement issue is addressed.
9. Addresses multiple, diverse controversies	Yes		Should be achievable.
10. Adaptable to all STs	Yes		Again, should be achievable.
11. Not based exclusively on 'yes'/'no' answers	Yes		No obvious preclusion here.
Notes: This model relies heavily on the engagement with external stakeholders. If this element can be suitably replaced then this modified version of Value Sensitive Design may			

meet the requirements of this research project.

Conclusions: The weighing of harms and benefits for different groups, and the identification of key values engaged, are two elements of Value Sensitive Design that may be transferable to other models.

Candidate: PRIVACY BY DESIGN (5.2.4)		**Category:** DESIGN-FOCUSSED APPROACHES	
Employment of other methods: Includes *stakeholder engagement*			
Criterion	**Can the candidate meet this criterion?**	**Could it be made to do so?**	**Comments**
1. For use before building commences	Yes		While this approach can be used to produce 'fixes' for privacy infringing STs, it aims to be *proactive* rather than *reactive* by preventing privacy infractions through the design process.
2. Results intended for senior use	Yes		No obvious preclusions here
3. Cannot involve external actors	Possibly	Yes	There is an element of stakeholder engagement within privacy by design. However, it should be possible to adhere to the 7 foundational principles of this approach within direct external engagement.
4. Produces design specifications, not policies	Yes		This is a very practical-orientated approach.
5. Output must add value	Yes		No obvious preclusions here.
6. Usable in any workplace	Yes		No obvious preclusions here
7. No prerequisite expertise required	Yes		Privacy by design seeks to capitalise on the technical expertise of the designers of STs by channelling these skills into creating privacy-sensitive products.
8. Minimal use-time required	Yes		More than many of the other models here, privacy by design feels like part of the design process itself rather than some 'added-on' process.
9. Addresses multiple, diverse controversies	No	No	Focusses primarily on privacy. This model is not designed to accommodate other common controversies, nor could it be modified to do so.
10. Adaptable to all	No	No	As per the criterion above, there is no relevance for this approach to the design of STs which do not infringe

STs			privacy.
11. Not based exclusively on 'yes'/'no' answers	Yes		No obvious preclusions here.

Notes: Privacy by design is concerned exclusively with privacy-related concerns. It is not applicable or modifiable to all STs and all commonalities of controversy.

Conclusions: While not an appropriate approach for extrapolation beyond the sphere of privacy-related concerns, there are two elements of privacy by design which should be adopted into future design tools within this project. The first is the proactive focus on producing mitigating effects during the design process rather than reacting to controversies as they occur. The second is the early identification of issues within the potential design of a ST which are then passed onto the designers themselves so they can utilise their technical expertise to produce solutions within their designs.

Candidate: COST-BENEFIT ANALYSIS (5.2.5)			**Category:** QUANTITATIVE ASSESSMENTS
Employment of other methods: N/A			
Criterion	**Can the candidate meet this criterion?**	**Could it be made to do so?**	**Comments**
1. For use before building commences	Yes		No obvious preclusions
2. Results intended for senior use	Yes		No obvious preclusions here.
3. Cannot involve external actors	Possibly	Yes	This depends on whether external actors need to be engaged with when determining which items are relevant for inclusion within any analysis and what weight to assign these.
4. Produces design specifications, not policies	Not necessarily	Indirectly	This approach assists in deciding between alternatives, however these alternative will already need to have been created. Cost-benefit analysis will not produce these alternatives to begin with.
5. Output must add value	Yes		Value will be added by indicating the 'best' choice between identified alternatives.
6. Usable in any workplace	Yes		No obvious preclusions here.
7. No prerequisite expertise required	Yes		No obvious preclusions here.

8. Minimal use-time required	Yes		No obvious preclusions here.
9. Addresses multiple, diverse controversies	Not all	No	Cost-benefit analysis will not be a suitable method for addressing all of the identified commonalities of controversy.
10. Adaptable to all STs	Yes		While cost-benefit analysis will be *usable* for all STs it will only be *of use* for those commonalities which lend themselves to a cost-benefit approach.
11. Not based exclusively on 'yes'/'no' answers	Yes		No obvious preclusions here.

Notes: An appropriate method for those questions which can be framed as a cost-benefit scenario. However it is not an appropriate tool for dealing with all commonalities in all STs.

Conclusions: A useful method when circumstances allow for such an approach, but probably best employed as part of a wider tool-kit of possible approaches.

Candidate: MULTI-CRITERIA DECISION MAKING (5.2.5)			**Category:** QUANTITATIVE ASSESSMENTS
Employment of other methods: N/A			
Criterion	**Can the candidate meet this criterion?**	**Could it be made to do so?**	**Comments**
1. For use before building commences	Yes		No obvious preclusions
2. Results intended for senior use	Yes		No obvious preclusions here.
3. Cannot involve external actors	Possibly	Yes	This depends on whether external actors need to be engaged with when determining which criteria are included, as well as what weight to assign these.
4. Produces design specifications, not policies	Not necessarily	Indirectly	This approach assists in deciding between alternatives, however these alternative will already need to have been created. Cost-benefit analysis will not produce these alternatives to begin with.
5. Output must add value	Yes		Value will be added by indicating the 'best' choice between identified alternatives.
6. Usable in any	Yes		No obvious preclusions here.

workplace			
7. No prerequisite expertise required	No	Possibly	Multi-criteria decision making is a developed field containing a number of different approaches and schools. Without previous expertise in using this method, the quality of any results will be questionable.
8. Minimal use-time required	Yes		No obvious preclusions here.
9. Addresses multiple, diverse controversies	Not all	No	Multi-criteria decision making will not be a suitable method for addressing all of the identified commonalities of controversy.
10. Adaptable to all STs	Yes		While multi-criteria decision making will be *usable* for all STs it will only be *of use* for those commonalities which lend themselves to this approach.
11. Not based exclusively on 'yes'/'no' answers	Yes		No obvious preclusions here.

Notes: An appropriate method for those questions which lend themselves to multi-criteria decision making. However it is not an appropriate tool for dealing with all commonalities in all STs.

Conclusions: A useful method when circumstances allow for such an approach, but probably best employed as part of a wider tool-kit of possible approaches.

Candidate: BRUCE SCHNEIER'S FIVE-STEP PROCESS (5.2.6)		**Category:** MISCELLANEOUS TESTS	
Employment of other methods: May include *stakeholder engagement* and *cost-benefit analysis*			
Criterion	**Can the candidate meet this criterion?**	**Could it be made to do so?**	**Comments**
1. For use before building commences	Possibly	Yes	It could be used during this point.
2. Results intended for senior use	Yes		No obvious preclusions here.
3. Cannot involve external actors	No	Doubtful	This all depends on what methods are required o obtain the necessary data for addressing the five requisite steps.

4. Produces design specifications, not policies	No	Possibly	This process is aimed towards the analysis and evaluation of STs, not at producing design specifications.
5. Output must add value	No	No	There is immense value created by this 5-step process in answering the question of whether or not a proposed ST is an appropriate response. This however is a step removed from adding value to the actual design of the ST itself.
6. Usable in any workplace	Possibly	Possibly	This depends on the resources of the designer, and their possession of the necessary data required answer the 5 steps.
7. No prerequisite expertise required	Possibly	Possibly	This all depends on the methods you need to employ to collect the necessary data for addressing the five steps of this process.
8. Minimal use-time required	Possibly	Possibly	As above, this will all depend on the time required to obtain the necessary data.
9. Addresses multiple, diverse controversies	No	No	This process is not designed to address specific controversies.
10. Adaptable to all STs	Yes		This process could be used on all STs, however the level of usefulness would vary considerably.
11. Not based exclusively on 'yes'/'no' answers	Yes		No obvious preclusions here.

Notes: This five-step process is excellent at analysing and evaluating security systems, STs, and processes. However, it is not tailored to address the design processes of building a ST as required under this research project.

Conclusions: This five-step process is not appropriate for this research project. However, there are elements of the five steps that have already been included within the identified commonalities of controversy.

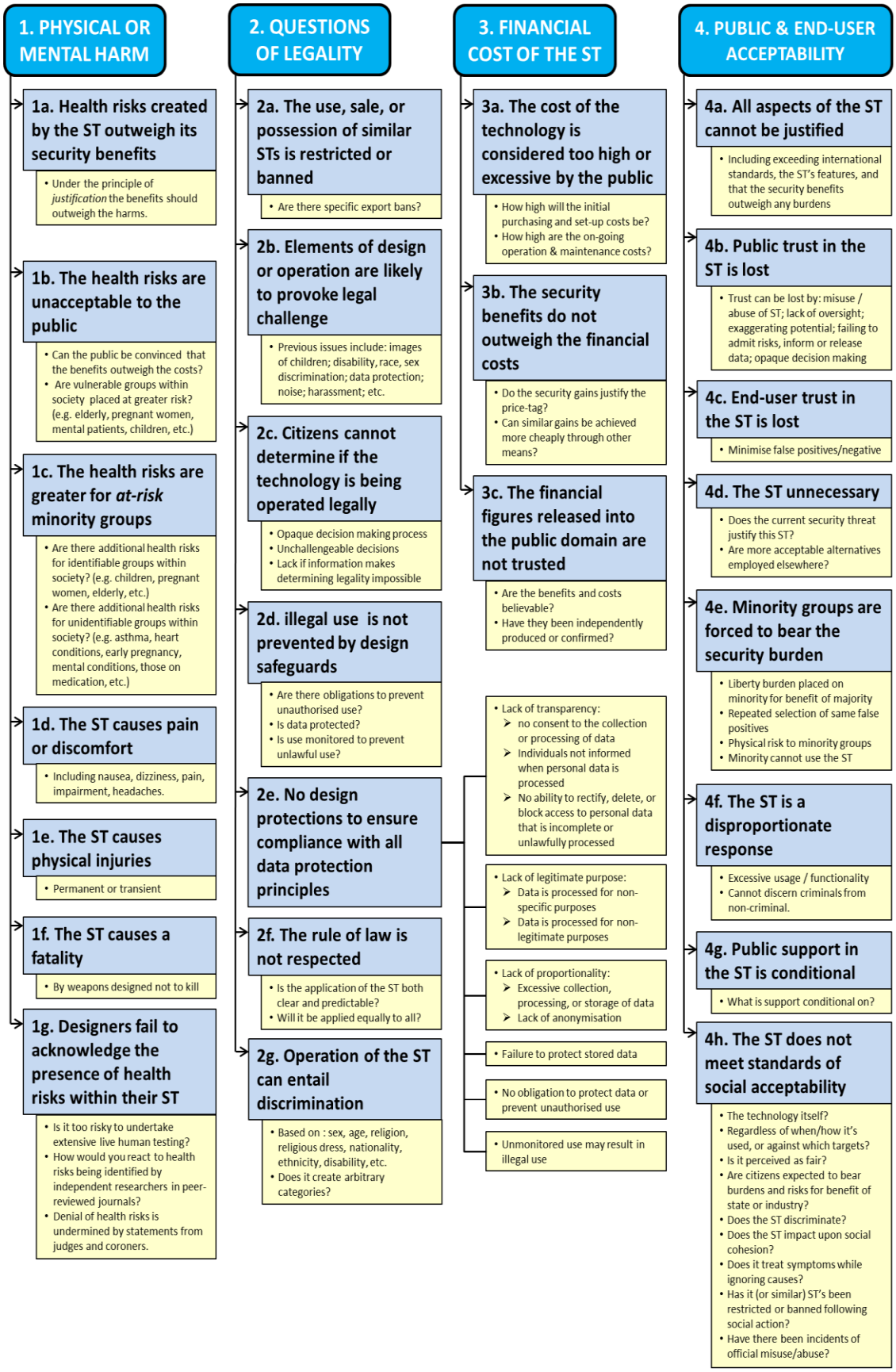
Candidate: ACLU'S NECESSARY AND DEFENSIBLE TEST			**Category:** MISCELLANEOUS TESTS
Employment of other methods: N/A			
Criterion	**Can the candidate meet this criterion?**	**Could it be made to do so?**	**Comments**
1. For use before	No	No	This test is not intended for use during the design of a

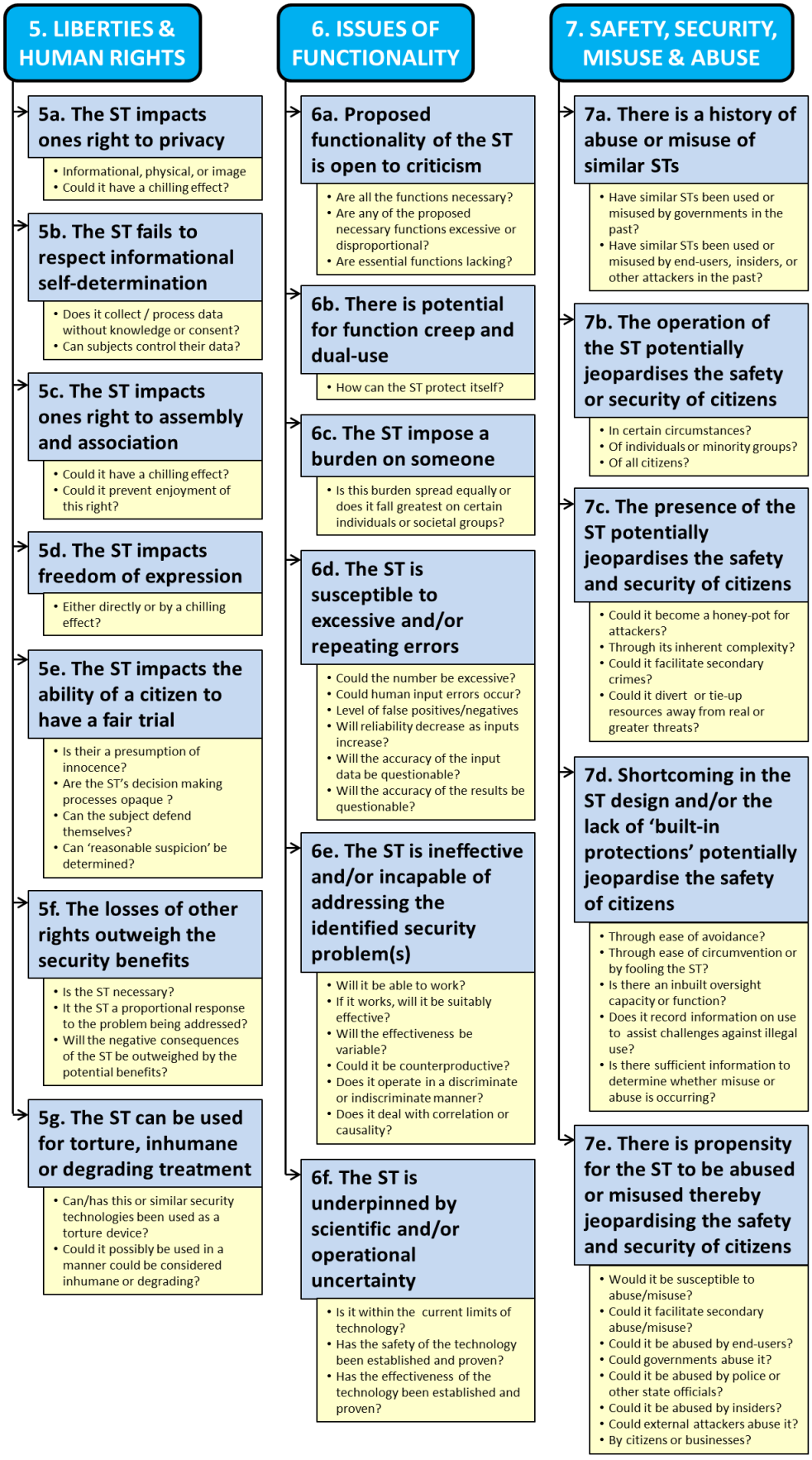
building commences			ST.
2. Results intended for senior use	Yes		No obvious preclusions here.
3. Cannot involve external actors	Yes		It could be conducted without external assistance.
4. Produces design specifications, not policies	No	No	It is aimed at policy makers, not designers.
5. Output must add value	No	No	Again, the value is for policy makers, not designers.
6. Usable in any workplace	Yes		No obvious preclusion here.
7. No prerequisite expertise required	Yes		No obvious preclusion here.
8. Minimal use-time required	Yes		No obvious preclusion here.
9. Addresses multiple, diverse controversies	No	No	This test is not designed for application to multiple controversies. Nevertheless, elements of value within the test itself can be found within a number of the previously identified commonalities of controversy.
10. Adaptable to all STs	Yes		This process could be used on all STs, however the level of usefulness would vary considerably.
11. Not based exclusively on 'yes'/'no' answers	Yes		No obvious preclusions here.

Notes: This test is aimed at policy makers and has little direct applicability to this project. However, there are elements of the questions asked which are of use.

Conclusions: Useful elements of this test have already been incorporated in the identified commonalities of controversy.

Appendix *K* Framework for Common Controversies within Security Technologies (FCC)




5. LIBERTIES & HUMAN RIGHTS
5a. The ST impacts ones right to privacy
• Informational, physical, or image
• Could it have a chilling effect?
5b. The ST fails to respect informational self-determination
• Does it collect / process data without knowledge or consent?
• Can subjects control their data?
5c. The ST impacts ones right to assembly and association
• Could it have a chilling effect?
• Could it prevent enjoyment of this right?
5d. The ST impacts freedom of expression
• Either directly or by a chilling effect?
5e. The ST impacts the ability of a citizen to have a fair trial
• Is their a presumption of innocence?
• Are the ST's decision making processes opaque ?
• Can the subject defend themselves?
• Can 'reasonable suspicion' be determined?
5f. The losses of other rights outweigh the security benefits
• Is the ST necessary?
• It the ST a proportional response to the problem being addressed?
• Will the negative consequences of the ST be outweighed by the potential benefits?
5g. The ST can be used for torture, inhumane or degrading treatment
• Can/has this or similar security technologies been used as a torture device?
• Could it possibly be used in a manner could be considered inhumane or degrading?
6. ISSUES OF FUNCTIONALITY
6a. Proposed functionality of the ST is open to criticism
• Are all the functions necessary?
• Are any of the proposed necessary functions excessive or disproportional?
• Are essential functions lacking?
6b. There is potential for function creep and dual-use
• How can the ST protect itself?
6c. The ST impose a burden on someone
• Is this burden spread equally or does it fall greatest on certain individuals or societal groups?
6d. The ST is susceptible to excessive and/or repeating errors
• Could the number be excessive?
• Could human input errors occur?
• Level of false positives/negatives
• Will reliability decrease as inputs increase?
• Will the accuracy of the input data be questionable?
• Will the accuracy of the results be questionable?
6e. The ST is ineffective and/or incapable of addressing the identified security problem(s)
• Will it be able to work?
• If it works, will it be suitably effective?
• Will the effectiveness be variable?
• Could it be counterproductive?
• Does it operate in a discriminate or indiscriminate manner?
• Does it deal with correlation or causality?
6f. The ST is underpinned by scientific and/or operational uncertainty
• Is it within the current limits of technology?
• Has the safety of the technology been established and proven?
• Has the effectiveness of the technology been established and proven?
7. SAFETY, SECURITY, MISUSE & ABUSE
7a. There is a history of abuse or misuse of similar STs
• Have similar STs been used or misused by governments in the past?
• Have similar STs been used or misused by end-users, insiders, or other attackers in the past?
7b. The operation of the ST potentially jeopardises the safety or security of citizens
• In certain circumstances?
• Of individuals or minority groups?
• Of all citizens?
7c. The presence of the ST potentially jeopardises the safety and security of citizens
• Could it become a honey-pot for attackers?
• Through its inherent complexity?
• Could it facilitate secondary crimes?
• Could it divert or tie-up resources away from real or greater threats?
7d. Shortcoming in the ST design and/or the lack of 'built-in protections' potentially jeopardise the safety of citizens
• Through ease of avoidance?
• Through ease of circumvention or by fooling the ST?
• Is there an inbuilt oversight capacity or function?
• Does it record information on use to assist challenges against illegal use?
• Is there sufficient information to determine whether misuse or abuse is occurring?
7e. There is propensity for the ST to be abused or misused thereby jeopardising the safety and security of citizens
• Would it be susceptible to abuse/misuse?
• Could it facilitate secondary abuse/misuse?
• Could it be abused by end-users?
• Could governments abuse it?
• Could it be abused by police or other state officials?
• Could it be abused by insiders?
• Could external attackers abuse it?
• By citizens or businesses?

Appendix *L* Designing for Socially Acceptable Security Technologies (DeSAST)

DESIGNING FOR SOCIALLY ACCEPTABLE SECURITY TECHNOLOGIES

Project Name:

_____

_____

_____

Completing Author(s):

_____

_____

Date: _____

Designing for Socially Acceptable Security Technologies **is a design tool created to assist the designers and developers of *security technologies** in anticipating, and thereby avoiding, potential design choices which are more likely to lead to social resistance when settling on the initial design specifications of a future security technology.**

Through identifying those design elements and choices which are more likely to result in citizens resisting or rejecting your security technology *before* they are incorporated into any initial product, you (the designer) afford yourself the best opportunity to create socially acceptable security technologies. By minimising controversy with socially responsive designs you will save money by preventing the need for future patches or modifications, as well as reducing the likelihood of social resistance which can lead to the regulation, restriction, or even the banning of a security technology.

The purpose of this design tool is *NOT* to tell you how to build your security technologies; rather it is to highlight where design choices can lead to future social resistance based on an evaluation of previous controversial security technologies. You can only consciously address a potential problem within a design after you have first identified an element as being problematic.

This design tool helps by asking you those questions which might not otherwise be asked. It remains your job to produce the best, most innovative and creative solutions when answering these questions; thereby adding value to your designs.

* A ***security technology*** is the product of an engineering endeavour which seeks to deter, prevent or detect crimes, and/or enhance the security of individuals, their property, or the state. This includes potentially lethal technologies; however they cannot be restricted to military use. Finally it encompasses all forms of engineering pursuit so long as what is produced has either a physical or digital presence. While a bike-lock, body-scanner, or security centric data-mining programme are all examples of security technologies, a policy, law, or general computer operating system will not suffice.

contents

If you choose to do so, to reduce the time taken to complete this design tool and to focus your attention on those sections most likely to possess the greatest relevance, you can use the following questions in italics to determine whether you should consider that section or not. For each question answered in the affirmative (using the information provided with each question to assist you in this determination) include that section when undertaking this design tool.

Caution: You are advised however not to disregard any section lightly. It may be that there are questions posed within a disregarded section whose relevance was not obvious to you when you chose to disregard that section.

1. PHYSICAL OR MENTAL HARM

Does the security technology (ST) possess the potential to cause physical or mental harm to the subject?

- **This harm does not need to be serious or life-threatening; for example any form of pain, discomfort, nausea, dizziness, physical or mental impairment will suffice.**
- **This harm does not need to be likely; statistically rare events of harm will suffice.**
- **The effects do not need to be permanent; transient harms such as short episodes of nausea, pain or dizziness will suffice.**
- **It is irrelevant whether this harm can be caused through either the use or misuse of the ST; statements to the effect that 'no harm will be caused if the end-user follows the designer's rules-for-use' have no bearing here.**
- **While the harm may be rare it must be plausible; i.e. using a body scanner as a weapon by crushing someone with one is not a plausible harm, however using a Taser-style baton as a club in close-quarters is a plausible harm.**

4. FINANCIAL COST OF THE ST

Given the financial cost of the ST, could questions be raised as to whether or not the ST represents a sound financial investment?

- **If you cannot demonstrate that the security benefits of your ST outweigh the financial outlay then questions remain over the prudence of this investment.**
- **If similar security gains can be achieved through other less expensive means, then this raises questions over the proposed ST.**
- **'Financial costs' do not just mean the purchase price; they also include any on-going operation, training, maintenance, calibration and upgrade costs, as well as any one-off charges (such as changes to building layouts, etc.).**
- **Failure to disclose financial costs to the public will leave them with unanswered questions.**

5. PUBLIC & END-USER ACCEPTABILITY

This Section incorporates a range of public and end-user concerns which have repeatedly had adverse effects on previous security technologies.

All security technologies should be assessed using this Section.

2. LIBERTIES & HUMAN RIGHTS

Does the ST possess the potential to infringe upon any of our other human rights when seeking to provide us with security?

- Human rights we are concerned with here include (but are not limited to) the following:
  - Privacy.
  - Informational self-determination.
  - Assembly.
  - Association.
  - Freedom of expression.
  - Right to a fair trial.
  - Freedom from torture or inhumane or degrading treatment.
- Remember that the provision of security does not automatically override the enjoyment of competing rights.

3. QUESTIONS OF LEGALITY

Does the potential exist for the legality of this ST to be questioned, challenged, or merely brought into doubt?

- This includes situations where: STs raise new issues of law where precedents have not been set; where the technology has not faced legal scrutiny; and where it is possible for somebody to evoke existing legislation to challenge the legality of this technology.
- STs which discriminate (or could be used in a discriminatory manner) should be considered.
- If the ST triggers data protection principles then this Section should be considered.
- If the potential exists for the ST to be used in an illegal manner then consider this section.
- If in the past this, or any similar, ST has been deemed illegal, subjected to bans or restrictions of use, then apply this section.

6. ISSUES OF FUNCTIONALITY

Could any of the functionality aspects relating to how the ST actually works (i.e., what it is does, what it doesn't do, what it's intended to do, and what it is capable of doing) be criticised or called into question?

- *Functionality* is defined here as all the things the ST can do and how it does them. It represents the culmination of all the individual *functions* (i.e. the individual components/ abilities/modes-of-use/ capabilities/etc.) built into the ST. Are you confident *all* the functions are necessary, proportional, and reliable?
- If aspects of the science or technology upon which the ST is built are new , unproven, or contain uncertainty then they will be called into question.
- It must be clear how the ST will operate, how effective it will be.
- If the ST suffers from errors, or is susceptible to function creep or dual-use, this will lead to questions over its functionality.

7. SAFETY, SECURITY, MISUSE & ABUSE

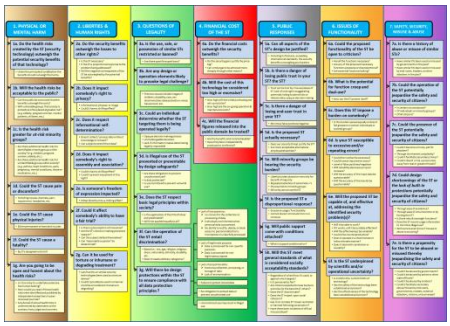
Could the ST be misused or abused thereby jeopardising the safety or security of citizens or their property? Additionally would it be possible for attackers to avoid or circumvent your ST?

- If there has been misuse or abuse of similar STs in the past then future misuse/abuse probably cannot be ruled out.
- If the safety or security of citizens depends on the ST being operated by end-users in accordance with your instructions then you cannot rule out the potential for misuse or abuse.
- Abuse or misuse of a ST can be by governments, police, state officials, insiders, end-users, businesses, external attackers, and citizens.
- Are you confident the ST will be impossible to avoid or circumvent, or that it will not become a *honey-pot* for attackers?

HOW TO USE THE DESIGN TOOL

step 1: Choose which of the 7 sections to include in your assessment of the ST

(option A) *Use all of the seven available sections as printed on the outside of the back cover* (see pages 436-437)



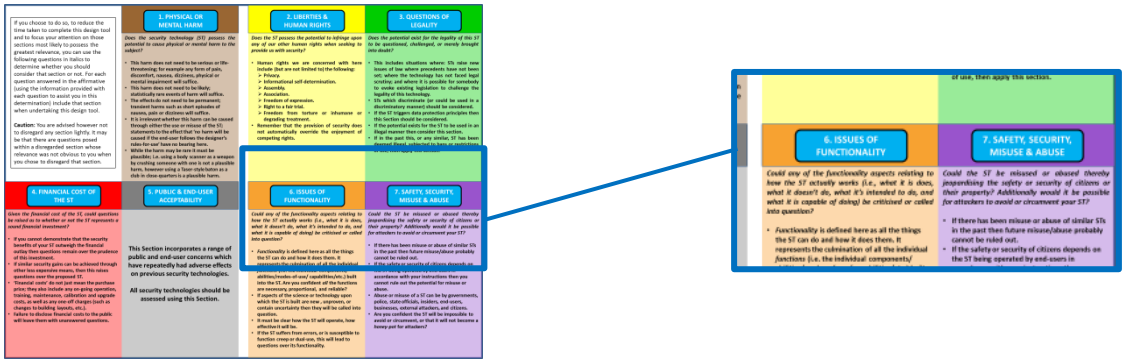
Benefit:

Ensures all available questions are addressed, including those which may have value to add despite not being obviously relevant at first glance.

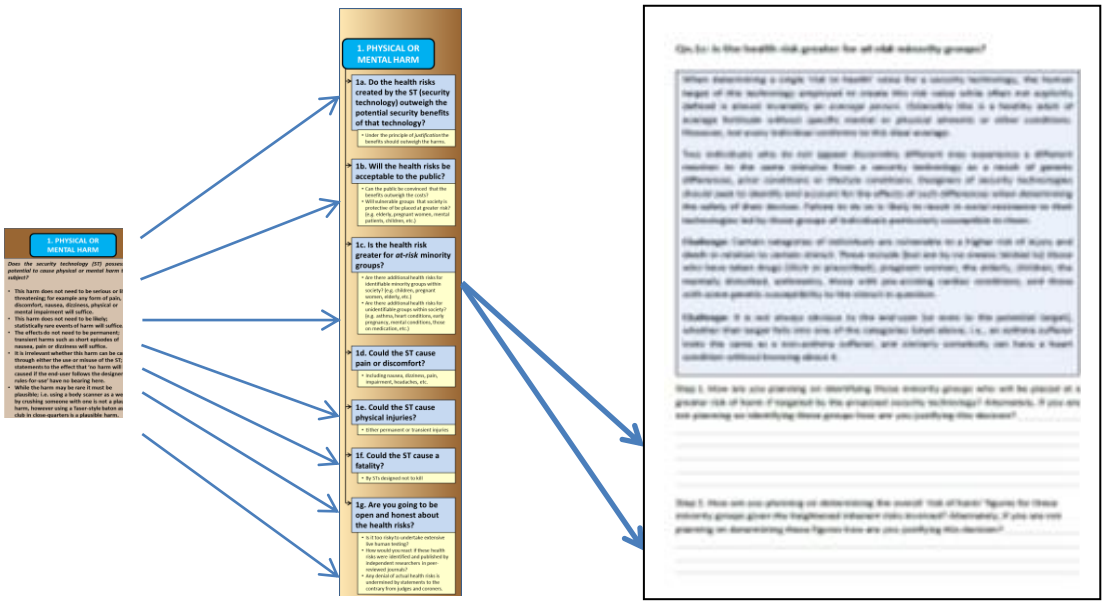
Drawback:

Increases both the length of time it will take to complete

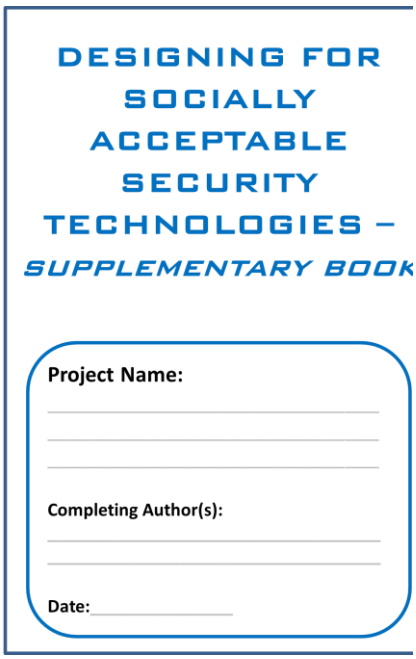
(option B) *Use the guide provided on the inside of the front cover to select which of the seven sections to include* (see pages 324-325)



step 2: For each included section(s) answer all of the questions contained therein, making sure (wherever possible and appropriate) to frame your answers as design opportunities



Use the accompanying ***Designing for Socially Acceptable Security Technologies – supplementary book*** *should you require extra space to answer the questions asked.*



To assist you in identifying and collating socially-responsive design opportunities:

(a) *Use the checklist found on the inside of the back cover to indicate both which sections were completed and the questions within the completed sections that are most likely to provide* ***design opportunities***, ***policy or public engagement opportunities***, *or represent* ***areas requiring future work***. (see pages 434-435)

(b) *Use the space provided in the section entitled* ***identified opportunities*** (page 109) *to bring together any* ***design opportunities***, ***policy or public engagement opportunities***, *or* ***areas requiring future work*** *you identified when completing this tool.*

important points to remember when answering the questions

- **Each *Section* comprises a number of different *questions*. These questions are then divided into a number of *steps* to assist you in answering the question.**
- **The point of each question (and its steps) is to get you to think about the design of your future security technology from a social, ethical perspective. The questions you are asked and the tasks you are set throughout this tool are designed to highlight those issues which have led to social controversy and resistance to security technologies in the past.**
- **You DO NOT have to answer all (or even any) of the steps which make up each question. Some will not be relevant to your particular security technology, while for others it is highly probable you will not possess the necessary information to do so. However for those questions which do apply it is important to remember that these are the questions a potential critic would ask when challenging the ethical and social credentials of your proposed security technology.**
- **The questions you are asked here are deliberately confrontational. They are designed to challenge the design of your proposed security technology and the reasoning behind this design. They are directed at the technology, not you, so please do not take offence at them!**
- **Just because you cannot answer any (or even all) of the questions contained within this design tool, or that you experience problems justifying your answers, this DOES NOT mean your proposed security technology will be socially controversial and/or unacceptable. However if a question is relevant and you find it hard or even impossible to answer, you may wish to consider the implications of such a situation carries with it.**

how to get the most out of this design tool *or* what this design tool *can* and *cannot* do for you!

- **This design tool CANNOT tell you how to build your security technology. You're the designer, that's your job!**

- **It CANNOT tell you the *best* design for your proposed security technology. Again, you're the designer, that's your job!**

- **It IS NOT a silver bullet for guaranteeing a socially acceptable security technology. But that is not to say we cannot learn lessons from the past.**

- **What this design tool CAN DO is to highlight the social and ethical design mistakes of the past so that *you*, the designer, can modify the design of *your* security technology so as to afford yourself the best opportunity to avoid repeating these mistakes. As to identifying what exactly the *best design solutions* are to these problems – you're the designer, that's your job!**

- **To this end, when answering the steps within each question, always try to identify design opportunities. How you answer the questions is completely up to you (there is no *right* or *wrong* way of doing so).**

- **If you find it useful you may wish to frame your answers as *future design requirements*. By collating these into a list once you have completed this design tool you will have produced a *practical set of design requirements* to be incorporated into the future design decisions of your security technology.**

section 1:

physical or mental harm

Qn.1a: Could the overall risks to health created by the security technology outweigh the potential security benefits of that technology?

A fundamental design principle when designing security technologies with the capacity to cause physical or mental harm to individuals is that these health risks should be outweighed by the opposing security benefits. This reflects the ethos behind the principle of *justification* within radiological protection whereby any exposure to radiation should do more harm than good.

There is no universal formula for determining whether security benefits outweigh health risks. This calculation will depend upon the characteristics of both the harm and the benefits; i.e., the type of harm (transient, permanent, death, etc.), frequency of harm, and nature of security benefit (national, group, individual, physical, property, etc.) will all need to be considered and converted into common values which can be weighed against each other.

In some cases a simple balancing of equivalent values (such as lives saved versus lives lost) may be possible. However when the values are not ostensibly equivalent (i.e. risk of lives saved versus injuries caused; risk of injuries caused versus property protected) then the user of this tool will need to produce an explicit formula for converting these variables to a common denominator. Additionally, if certain types of harm and/or certain security benefits possess greater weight within the public psyche, then the developer is advised to apply a weighting factor to those values when comparing harms and benefits.

Challenge: The fact a security technology entails a very low risk to health does not automatically mean the benefits outweigh the risks. Also be aware of situations where a security technology is used very often while instances of attacks are very rare.

- Airport backscatter whole body scanners attracted health concerns despite delivering an incredibly small dose (0.02 μSv) of ionising radiation to those screened with an estimated risk of death from a single scan of one in one-billion. Given that more than 5 billion passengers fly each year, even this very small dose would statistically result in deaths.

Challenge: In a hostage (or similar life-threatening) scenario where a security technology which is deployed to save lives results in a high death or injury count, this technology may prove controversial and be hard to justify even if more than 50% of people are spared death or injury as a result of its deployment.

- The use of an undeclared aerosolised anaesthetic agent to help end the terrorist siege of Dubrovka Theater, Russia, resulted in the death of conservatively 129 out of 850 hostages. This resulted in considerable controversy despite the majority of hostages being saved.

Challenge: Despite a security technology possessing the capacity to cause harm, unless data already exists you will need to collect sufficient data from live experimentation. As a potential design 'rule-of-thumb' if it is too dangerous to test an ostensibly non-lethal security technology on human subjects because of the risk of death or serious injury, then justifying the deployment of this technology will be extremely difficult.

- Less lethal weapons designed to subdue or disperse individuals, such as Tasers, baton-rounds, gases, microwave technologies, etc., can all be fatal. Nevertheless testing on human subjects still needs to occur before deployment can be justified.

Step 1. What are the potential harms and the anticipated rates of these harms?

Step 2. Apply a weighting factor to these harms if applicable.

Step 3. What are the anticipated security benefits and their rates?

Step 4. Are the 'rates of harms' and 'security benefits' common values which can be compared? If 'yes' compare harms and benefits. If 'no' first convert harms and/or benefits into a common value which can be compared before comparing harms and benefits.

Step 4. Do the security benefits outweigh the harms? If 'yes' what concerns remain regarding this calculation? If 'no' how can development and deployment of this technology be justified

Qn.1b: Will the health risks be acceptable to the public?

Assume a starting position whereby according to statistical, experimental, and/or modelled data the security benefits of a proposed security technology outweigh the health risks associated with its use. This fact, while of fundamental importance, may not in itself be enough to ensure social acceptability on the issue of health, and it is unlikely a security technology which poses health risks (even if minimal) will survive without modification or restrictive legislation mid to long-term if it is not considered acceptable by the public. Further actions may be required to maximise the likelihood that the public will accept these health risks, however such an outcome can never be guaranteed.

Challenge: Failure to convince the public that the health risks are not just minimal but also justified will potentially undermine public support. This may require the release of accurate data, public engagement, and assessment of the technology by independent experts who will enjoy public trust.

- Backscatter whole body scanners were the subject of intense public and governmental debate regarding the associated health risks despite these risks being incredibly low. This delayed the mainstream introduction of this technology to the benefit of the alternative millimetre wave scanner technology despite it being less capable of detecting concealed objects.

Challenge: If certain vulnerable minority sub-groups within society are at greater risk of harm by the security technology than the rest of the population, this fact may have an negative bearing on the overall acceptability to society of this technology; especially if this wider society is particularly protective of these sub-groups. Examples here potentially include pregnant women, children, the elderly, disabled, mental patients, etc.

- Use of less lethal weapons, such as Tasers and pepper-sprays, on individuals from these vulnerable groups has in the past resulted in a negative societal response.

Step 1. Have you ascertained the level of health risk?_____ Has this level been independently verified, and if so by whom?_____

_____

_____

_____

If this level has not been independently verified, why was this decision taken?_____

_____

_____

_____

_____

Step 2. Are you permitted to discuss the proposed security technology with the general public?_____ If 'yes' complete Step 3; if 'no' complete step 4.

Step 3. What is your plan for engaging with the public to ascertain the acceptability of the proposed security technology? Alternatively, if you are permitted to discuss this security technology with the public but are not planning on doing so, how are you justifying this course of action?_____

Step 4. How are you planning on ascertaining the acceptability of the proposed security technology given you are not permitted to directly discuss it with the public?_____

Note on Step 4: Viable methods for ascertaining the acceptability of a security technology which you are not permitted to discuss directly with the public may include:

- Creating a number of theoretical proxy technologies which each contain different elements of the *genuine* security technology and assessing these with focus groups.
- 'In house' workshops with individuals not directly involved on the project but with sufficient clearance, focussing on their reactions to this technology.
- Assessment of public reactions to similar technologies that already exist.

Step 5. If *health risks associated with the proposed security technology* are identified as a cause for concern in Steps 3 or 4, ask yourself the following questions:

- What design modifications, if any, could mitigate these concerns?
- Are there situations or circumstances where the need to use this proposed technology outweighs the social concerns over the health risks thereby making this technology acceptable?

Qn.1c: Is the health risk greater for *at-risk* minority groups?

When determining a single 'risk to health' value for a security technology, the human target used to create this risk value (while often not explicitly defined) is almost invariably an *average person*. Ostensibly this is a healthy adult of average fortitude without specific mental or physical ailments or other conditions. However, not every individual conforms to this ideal average.

Two individuals who do not appear different may experience different reactions to the same stimulus from a security technology as a result of genetic differences, prior conditions, or lifestyle conditions. Designers of security technologies should seek to identify and account for the effects of such differences when determining the safety of their devices. Failure to do so may result in social resistance to their technologies led by those groups of individuals particularly susceptible to them.

Challenge: Certain categories of individuals are vulnerable to a higher risk of injury and death in relation to certain stimuli. These include (but are by no means limited to) those who have taken drugs (illicit or prescribed), pregnant women, the elderly, children, the mentally disturbed, asthmatics, those with pre-existing cardiac conditions, and those with genetic-based susceptibility to the stimuli in question.

Challenge: It is not always obvious to the end-user (or even to the potential target), whether that target falls into one of the categories listed above; i.e., an asthma sufferer looks the same as a non-asthma sufferer, and somebody can have a heart condition without knowing about it.

Step 1. How are you planning on identifying those minority groups who will be placed at a greater risk of harm if targeted by the proposed security technology? Alternately, if you are not planning on identifying these groups how are you justifying this decision?

Step 2. How are you planning on determining the overall 'risk of harm' figures for these minority groups given the heightened inherent risks involved? Alternately, if you are not planning on determining these figures how are you justifying this decision?

Step 3. List off all of the 'at-risk' groups then, one at a time, insert each group into the following paragraph: "A *(insert group name here; e.g. pregnant woman, asthmatic, child, etc.)* suffers a serious injury related to their risk-condition as the direct result of being targeted by the security technology. The use of the technology and subsequent injury was witnessed by members of the public and recorded on a phone with the footage now being replayed on national television". For each 'at-risk' group describe what you believe to be the *best-case* and *worst-case* public reactions to this footage.

Step 4. How can you build into this proposed security technology methods for protecting these 'at risk' groups?

Note on Step 4: The challenge is to build physical protections into the technology itself so you are not left relying solely on the perception, judgement, and training of the end-user. How these would/could work depends of course on the nature of the technology itself, the anticipated targets, and the situations in which it is expected to be used. Approaches to addressing this challenge might include:

- Methods whereby the security technology itself identifies the target as inappropriate given the risk, potentially with the ability for the operator to override this decision if they believe circumstances warrant this call.
- Methods whereby the potential target is afforded the opportunity to identify themselves before deployment of the security technology.
- Alternative settings for the security technology whereby the risk to this group is appropriately managed.

Qn.1d: Could the security technology cause pain or discomfort?

Security technologies with the capacity to cause pain or some lower level of discomfort (e.g. dizziness, nausea, headaches, physical or cognitive incapacity/impairment, etc.) always run the risk of evoking emotive societal responses. Images on mainstream media and social-networking sites of non-violent protestors, the elderly, defenceless individuals, and/or children being subjected to such technologies have in the past produced outpourings of social anger. That being said, the fact a security technology is designed to cause pain or discomfort to achieve some legitimate goal in no way *automatically illegitimatises* that technology. While less lethal weapons and crowd control technologies, including Tasers, teargas, and long range acoustic devices have all been the source of social controversy in certain instances, in others they have been successfully deployed without social resistance. It is this contradiction in responses towards what is often the same technology based upon the context and manner of its use which provides both challenges and opportunities for clever developers.

Challenge: Designing a technology which can cause pain or discomfort in a socially acceptable manner; i.e. which can as far as is possible - discriminate between offenders and non-offenders, is proportionate and serves a legitimate aim, is not open to abuse misuse or overuse, has effects which can be controlled and which end as soon as an offender leaves the area or stops their illegal actions.

In the past:

- Mosquitos have been criticised for not discriminating between offenders and non-offenders, preventing young people enjoying their right to assembly, for being used by shopkeepers to control public spaces which they do not own or have a right to control, and causing distress to babies and very young children who cannot communicate the source or nature of their pain to their parents who can't themselves hear the noise being emitted so are unaware of its use.
- Use of less lethal weapons such as Tasers and pepper sprays by police as either a form of punishment or against non-violent protestors or unarmed citizens who are arguing with them, talking back, walking away, or being discourteous undermines public support for these weapons and the people who use them.

Step 1. How is the level of pain/discomfort produced proportionate to the intended security benefit?_____

_____

Step 2. Can the end-user control the level of pain produced?_____ If 'yes' what protection measure prevent misuse here?_____

_____

Step 3. Does the proposed security technology target a single individual or is it designed to affect a physical area?_____ If it affects an area, what controls are you going to include in your design so the end-user can modify the range/shape of the area affected, and the effect on individuals within the affected area?

Note on Step 3: 'Area effect' security technologies are often blunt instruments, lacking the ability to moderate their effects depending on circumstances or to discriminate between legitimate and illegitimate targets. This can lead to negative social responses. The ability to moderate the effect of the technology, to swiftly change the shape of the physical area affected, or (as the Holy Grail in this area) to be able to only effect 'wrongdoers' while not affecting innocent bystanders and security personal within the range of the technology may help minimise the probability of negative responses.

Step 4. Will the pain/discomfort being applied to the target cease **automatically** if they leave the zone-of-effect or stop the illegal action(s), or does this rely upon the end-user either taking or stopping some action (i.e., like taking their finger off a trigger, etc.) before the pain-inducing effects will stop?

If it is the latter, how can misuse or overuse by the end-user be prevented, especially in situations where the end-user is emotional (i.e., scared, angry, hyped-up, stressed, etc.) and hence may not be acting rationally or displaying sufficient self-control?

Step 5. How long until the pain/discomfort effects of the security technology cease for the target?
Can this timeframe be reduced, and if so how?

Step 6. To help minimise abuse, misuse, overuse, or illegitimate use of the proposed security technology, how can its use be monitored by the technology itself?

Step 7. How can the presence and zone-of-operation of the security technology be made obvious to everybody both within and without the affected area?

Qn.1e: Could the security technology cause physical injuries?

This section does not apply to security technologies which are designed to kill; for these technologies move on to Qn.1f.

When discussing 'physical injuries' we have moved beyond mere pain or discomfort. Injuries we are concerned with here include puncture wounds, broken bones, internal organ damage, burns, damage to mobility or senses like sight, hearing, etc. beyond a transient loss of control, brain damage, severe bruising, muscle damage, nerve damage, etc. Within security technologies these types of injuries will most commonly be caused by less lethal weapons (various forms of baton rounds, Tasers, water cannons, etc.) however for our purposes it would also include non-fatal cancers resulting from those technologies which employ ionising radiation.

Security technologies which are designed with the specific, deliberate, and ***sole*** purpose of causing the physical injuries detailed above will not enjoy widespread public support, will face active public resistance, and the use of these will be subject to legal challenge. This is because any security technology designed with the sole purpose of causing such injuries but without killing the target is essentially an instrument of torture. It is also highly likely any company developing such technologies would risk reputational damage.

However, security technologies designed for specific security purposes (protecting citizens or property, regaining order during a riot, apprehending offenders, etc.) but whose deployment runs the risk of unintentional physical injuries can enjoy public legitimacy and support. Though gaining and maintaining this support will be challenging, especially if the benefits are not obvious or in the face of instances of abuse.

Challenge: The need to minimise both the risk and severity of unintentional physical injuries resulting from the legitimate use of a proposed security technology whilst maintaining the effectiveness of that technology, is a challenge which designers must address.

Challenge: Designers will need accurate data on the injury risks of their technology if they are to justify claims that the security benefits outweigh the health risks. So to determine the *risk of injury* from deploying the proposed security technology, a sufficient amount of 'live' testing will be required for accurate data to be obtained. While simulation may play a part here, if the risk of live testing is too dangerous then the technology itself is probably too dangerous to deploy.

Step 1. How can you design the proposed security technology so as to reduce the likelihood of injuries occurring while maintaining effectiveness?

Step 2. How can you design the proposed security technology so as to minimise the severity of those injuries which do occur?

Step 3. Accurate data on the expected injury risks will be required to quantify the risks involved. If accurate data is not available outline your plans for creating this data.

Step 4. Would it be too dangerous to conduct live-testing of your proposed security technology? If 'yes' how are you justifying the deployment of your technology given the obvious dangers involved?

If 'no' set out in detail your plans for the live testing of your proposed security technology?

Qn.1f: Could the security technology cause a fatality?

It is possible to divide security technologies into three categories: (1) technologies specifically designed to kill, (2) technologies designed not to kill but possessing the capacity to do so, and (3) technologies which cannot kill regardless of design.

Those in category 1 are not restricted to firearms. They include corporal punishment technologies (e.g. electric chairs, lethal injection drug-cocktails and administering devices), lethal perimeter security technologies (electrified fences, lethal automated response systems), etc. As these technologies are overtly designed to be lethal, it is not the fact they subsequently cause a fatality when deployed (effectively operating exactly as intended) that will be the source of controversy. Rather it is either; the fact they exist at all within a in society, the operational circumstances surrounding individual instances of use, characteristics of the target killed, or the manner in which they achieve their aims, which will create the negative social reactions.

Those in category 3 are not relevant to the current Question.

The main concern of this Section is those in category 2. Predominantly this includes a number of less lethal weapons such as Tasers, the various baton rounds, tear gas, water cannons, etc. It will also include technologies which subject the target to a dose of ionising radiation.

Negative social reactions can arise here in a number of different ways: (a) If subjecting one's self to the technology is a mandatory prerequisite for accessing a necessary service (such as an airport screening system), then any risk of death for the individual will be controversial, even if that risk is very small and outweighed by the security benefits. (b) If the technology is marketed, promoted, and deployed as an alternative to lethal force and it then kills someone, this can also easily undermine credibility in the technology and lead the public to question its continued use.

Step 1. How can you design the proposed security technology so as to reduce the likelihood of death occurring while maintaining effectiveness?

Step 2. Accurate data on the expected risk of death will be required to quantify the risk involved. If accurate data is not available outline your plans for creating this

data.

Step 3. Would it be too dangerous to conduct live-testing of your proposed security technology? If 'yes' how are you justifying the deployment of your technology given the dangers involved?

If 'no' set out in detail your plans for the live testing of your proposed security technology?

Step 4. Given the risk of death involved, would you be prepared to have the technology either tested on, or deployed in the field against, either you and/or members of your family? If 'no' how can you justify the proposed design of this technology?

Step 5. Could misuse of the proposed security technology increase the likelihood of a death occurring? If 'yes' how can you design your technology to prevent/minimise misuse?

Qn.1g: Are you going to be open and honest about the health risks?

Failing to be open about the health risks associated with a security technology is a risky ploy which can only produce short-term benefits (such as embedding a technology into the environment and/or increasing market share). However, these short-term benefits may be outweighed by any medium to long-term social backlash in response to the initial lack of disclosure, for it is impossible to prevent these health risks being made public.

The health risks of a security technology are impossible to hide and will be made public through a variety of avenues. Curious independent researchers will publish information via peer reviewed journal articles. Individuals who are harmed by the security technology will seek the disclosure of previously known information from the developer in court. Government inquiries/hearings will draw out information. Pressure groups and NGOs will challenge the safety of security technologies. Regulators will seek information.

Despite the fact it is impossible to hide health risks, developers continue to adopt this approach.

- During the early stages of development, the manufacturers of whole body scanners adopted the line that the low intensity X-rays did not penetrate the subject and were merely reflected off their skin. Eventually they corrected this error.
- The electroshock weapon manufacturer Taser International repeatedly and vigorously denied their weapons could produce any adverse health effects, suing a researcher for publishing peer-reviewed critical scientific evidence and a medical examiner for listing the Taser as a cause of death on a death certificate. They also created a refuted mental state *excited delirium* as an alternative to those deaths where a Taser was deployed. Both judges and Taser International have now explicitly acknowledged their health risks.

This drip-feeding of information and/or the reversal of positions on the health risks of a security technology may undermine the credibility and trustworthiness of the developer in the eyes of the public. The volume of controversy surrounding the technology will also potentially be increased as now there are two stories for the press; the health risks inherent to the technology as well as the failure to disclose these health risks by governments and manufacturers.

What is surprising about refusing to acknowledge the health risks of a security technology is that ultimately this is an unnecessary tactic. There are numerous security technologies which pose a risk to the safety of citizens which are still considered acceptable within society. As people are willing to accept a level of health risk, developers can legitimately frame the presentation of their technology by asking "*Is it as safe as, or safer than, the alternatives?*"

Finally, the failure (by yourself or others) to test for health risks does not allow you to make the claim "*My product is safe*" as this absence of evidence on possible health risks is not evidence of the absence of health risks.

Step 1. Are you able to inform the public about the health risks of the security technology?_____ If 'yes' move on to step 2. If 'no' explain here why not:

Step 2. Assuming you are able to inform the public about the health risks of the security technology, are you going to?_____ If 'yes' complete step 3 only. If 'no' move on to step 4.

Step 3. What is you plan for dissemination of these risks?

Note on Step 3: When producing this plan consider the following questions: Where will you disseminate this information? How detailed will it be? Will you engage directly with the public before and/or after the technology is released? How will this public engagement occur and what will it entail? How and when will you engage with those citizen groups, NGOs, and issue groups which traditionally oppose such security technologies? etc.

Step 4. Are you going to actively deny the health risks?_____ If 'no' move on to step 5. If 'yes' explain here what benefits you expect to gain from this course of action, and set out how these benefits outweigh the risks of the future disclosure of the health risks? (Alternatively if you are too worried about the legal implications of filling out this section, perhaps you may wish to reflect upon your decision in this step)

Step 5. What is your plan for responding to questions/accusations when the facts about the health risks become public?

Note on Step 5: Consider here both the nature of your response (litigation, continued denial, silence, acceptance, information release, etc.) and the source of the public disclosure (former/current employees, government leak or official release, courts, independent researchers, peer reviewed journals, etc.). How will these affect the continued operation and acceptance of the technology, as well as the reputation of you the designer and you clients?

section 2:

liberties & human rights

Qn.2a: Could the security technology impact someone's right to privacy?

An individual's right to privacy is recognised under various international conventions, treaties, national legislation, and constitutions. However, this right is rarely absolute as there are numerous legitimate reasons for encroaching upon an individual's privacy, including; national security, preventing and investigating crimes, and the provision of public services to name but a few.

While there may be security gains to be had by restricting privacy, the infringement of someone's privacy by a security technology is probably the single hot-topic issue which attracts the most attention by the public, politicians, and the media. Previously largely restricted to physical privacy (i.e. the collection of physical images by CCTV, physical correspondence, or voice recordings by tapping), privacy now also includes informational privacy (i.e. our internet usage). It has become a focal point for many groups of citizens opposed to the expansions of security technologies, whether online or in public spaces.

What is unfortunate is that privacy and security are often presented as diametrically opposing forces (i.e. decreasing privacy increases security, and increasing privacy decreases security) through the use of the commonly seen *balancing metaphor*. This simplistic interpretation ignores the fact that increasing someone's privacy can simultaneously increase their security, and *vice versa* by decreasing their privacy we can decrease their security. There is also the issue of both targeted and untargeted privacy-infringing surveillance having a chilling effect on society when it comes to exercising other rights, such as assembly, lawful protest, expression, etc.

Challenge: The challenge for designers is to develop security technologies which:

- Protect privacy through more targeted approaches;
- Increase our security *and* privacy simultaneously;
- Do not make people less likely to exercise other legitimate rights out of fear of government surveillance;
- Are socially acceptable in an environment where privacy infringing technologies *will be scrutinised* by activist organisations and a sceptical public;
- Are able to justify any privacy infringement as necessary, proportionate, and productive (i.e. able to show real security gains for any privacy loss).
- Respect citizen's presumption of innocence.

Step 1. What new security will the proposed technology provide us with?_____

What (if any) privacy will we lose or gain because of the proposed security technology?_____

Justify any changes in security and privacy resulting from your technology_____

(see also Qn.5f)

Note on Step 1: To justify the proposed security technology we must either; (a) gain security or privacy without losing the other, or (b) gain a specific amount of security which warrants the loss of a specific amount of privacy.

If the situation is 'b' then you will need to convert security and privacy into common values so as to meaningfully balance these two concepts if you are to justify the security/privacy effects of your technology. There exists no generally accepted meta-rule or formula for doing this so your justification will unavoidably be open to questioning.

Step 2. What privacy is lost by people who are not engaged in criminal activities but are subjected to the proposed security technology (i.e. what is the privacy cost to innocent people)?

Step 3. What privacy will citizens be left with if the proposed security technology is successfully introduced?

Step 4. What are the possible negative security implications of this privacy intrusion?

Note on Step 4: Consider; insider threats, misuse/abuse by officials, theft of collected data, how this information could be used to harm a particular person, etc.

Step 5. How can you minimise both any impact on privacy and any risk of negative security implications through the design of your security technology?

Note on Step 5: Design ideas here may include: anonymisation of data/images etc., technologies which only record events of interest, limiting the time data is held, etc.

Step 6. If you were personally being monitored by your proposed security technology, would you be as likely to engage in the following legal activities: having an affair?_____; writing a blog criticising your employer/the government/police?_____; sending/receiving legal documents to/from your lawyer?_____; organising or participating in a protest rally?_____; purchasing, downloading, and/or viewing pornography?_____; seeking online advice on abortions/ medical conditions/assisted suicide?_____; supporting/sending money to Wikileaks?_____. If you answered 'yes' to any of these your technology is having a chilling effect. How can you minimise this effect through your design?

Qn.2b: Does your proposed security technology affect the ability of citizens to exercise informational self-determination?

Informational self-determination is the right of individuals to decide; what information about themselves should be communicated to others, under what circumstances this should occur, and what should happen to this information if it is communicated. It is a concept which has been attracting considerable attention throughout the EU over the past few years and encompasses ideas of privacy and consent.

However, when seeking to provide security and/or tackle crime there is the obvious need to collect information on individuals which may override the ideal of informational self-determination. If somebody engaged in illegal activities could legitimately prevent police accessing information about these activities by hiding behind this principle, the implications for society would be severe.

That being said, the provision of security and the prevention of crime are not automatic trump cards when it comes to competing rights. When this fact is either forgotten or not respected by governments and/or the developers of security technologies we see incidents of social resistance. The collection, storage, processing, and forwarding of personal data must be in response to genuine crime/security problems, necessary for investigative work, and proportional to the intervention.

Situations where controversies have arisen in the past include incidents where; collected data was not securely stored, collected data was mined for purposes other than which it was collected thus resulting in criminal *fishing expeditions*, the affected party was not able to inspect the stored data and change erroneous information, data was held indefinitely regardless of the guilt/innocence of the individuals, and where data was transferred to third parties/countries either illegally or where the receiver does not have effective data protection systems.

Challenge: To develop security technologies which ***simultaneously***:

- Respect informational self-determination including the ability to control, inspect, and amend held personal data.
- Facilitate police in their investigations of known or suspected offences without alerting suspects.
- Prevent fishing expeditions and unauthorised/illegal access of the data.
- Ensure secure data transference, handling, and storage.

Step 1. Will your security technology be able to identify the purpose(s) for which personal data was collected so that it can subsequently prevent the processing of this data for other purposes?_____ If ‘no’ why is this the case. Alternately if ‘yes’ how will

you achieve this?

Step 2. Will your security technology be able to verify that any search/processing of the collected data is based on appropriate reasonable suspicion formed from other evidence so as to prevent fishing expeditions? If 'no' why is this the case. Alternately if 'yes' how will you achieve this?

Step 3. If personal data is held on a person, how will they know this is the case?

Will they be afforded the ability to view this data and to correct/challenge any errors, and if so how will this process be made clear to them?

Will the security technology include a capability allowing for either *real-time* corrections of false data by end-users in the field (such as security screeners) or regular updating of the data in response to feedback from end-users or the results of other security technologies, thereby limiting false positives? If 'no' why is this the case. Alternately if 'yes' how will you achieve this?

Step 4. How will your proposed security technology control access to the data to both keep it secure and to prevent unauthorised access by insiders?

Step 5. Will your proposed security technology be able to simply, quickly, and effectively delete data that is no longer needed, no longer relevant, or where it is no longer appropriate that the data be held? If 'no' why is this the case. Alternately if 'yes' how will you achieve this?

Qn.2c: Could your technology influence the enjoyment of an individual's rights to assembly or association?

Under the European Convention on Human Rights citizens have the right to freely associate with other people and to freely assemble together. However, as with nearly all such rights these are not absolute. Restrictions may be placed on them if they are prescribed by law *and* necessary in our society to; protect the interests of national security or public safety, for the prevention of crime or disorder, to protect health and morals, and to protect the rights and freedoms of others.

There are two main ways in which security technologies have proven controversial with the public in the past. Firstly by actively preventing citizens from exercising these rights, and secondly via a *chilling effect* whereby citizens are not willing or less willing to exercise their rights out of fear their actions will be monitored and/or they will face some future repercussion(s) if they do exercise their rights.

Security technologies in this first category include mosquitoes which force young people out of public spaces where they congregate, less lethal weapons being used to disperse protestors, and technologies which control access to communication services (i.e., phones, internet, etc.). Security technologies in the second category include mosquitoes, less lethal weapons, CCTV, online monitoring and data mining services, etc.

Challenge: This brings us to the challenge for developers of security technologies which impact upon someone's rights to assembly and association: *How do you develop a technology for restricting these rights which under normal circumstances can be exercised freely by all citizens without lawful impediment, such that:*

- Your security technology cannot be used in situations where its use is not lawful *or* its use is not necessary to achieve the legitimate goals specified earlier.
- The presence of your technology does not have a chilling effect on citizens who wish to exercise their rights to assembly or association.
- Your security technology is effective in preventing association or assembly when it is both lawful and necessary to do so.

Step 1. Identify the circumstances under which it will be ***lawful*** to use your proposed security technology when it is having the effect of infringing someone's right to assembly or association?_____

_____

_____*If you cannot identify any such circumstances then chances are it will be unlawful to use your proposed technology.*

Step 2. Identify scenarios whereby it will be ***unlawful*** to use your proposed security technology when it is having the effect of infringing someone's right to assembly or association?

What protections could you possibly build into your security technology to either prevent or minimise the likelihood of such unlawful use occurring?

If your technology is open to unlawful use and there are either no technical protections available (or there are but you are not going to build-in these protections), justify why your technology should be built.

Step 3. Explain how your technology will achieve any or all of the following:

- protect national security or public safety;
- prevent crime or disorder;
- protect health and moral;
- protect the rights and freedoms of others.

Why is your proposed technology *necessary* to achieve these legitimate aims?

Note on Step 3: If there is another existing technology/approach which can achieve the same goals as your technology with less social impact then your technology probably cannot qualify as a necessity.

Step 4. If you were personally being monitored by (or subjected to) your proposed security technology, would you be as likely to, or worried about, engaging in the following legal activities (which for these purposes you wish to engage in): having an affair?_____; applying for a role at a competitor's firm?_____; attending a meeting of the BNP?_____; organising or participating in a protest rally?_____; attending an alcoholics anonymous meeting/drug-counselling session/mental health support group?_____; attending a religious sect meeting?_____. If you answered 'yes' to any of these your technology is having a chilling effect. How can you minimise this effect through your design?

Qn.2d: Could the security technology restrict somebody's freedom of expression?

Everyone has the right to freedom of expression; the right to hold opinions and to receive and impart information and ideas without interference by public authorities. Therefore we have the right to say what we want, write what we want, and publically demonstrate so as to publicise these views. However, this right is not absolute. Where necessary it may be restricted in the interests of national security, public safety, and the prevention of crime and disorder, amongst other reasons.

Security technologies can and are used to directly infringe this right, largely in the *digital* world; for example the blocking of certain websites by Internet Service Providers and the monitoring of internet users who visit certain sites or seek to access certain information. The presence of such technologies can also have a chilling effect, making it less likely people will exercise their right of expression. This restriction/monitoring can either be targeted at individuals or a blanket approach covering entire nations. In the physical world security technologies such as CCTV cameras and less lethal weapons can make people wary of exercising their right of expression.

From a social perspective, security technologies which seek to restrict somebody's freedom of expression or to monitor their actions are often a source of mistrust and anger. At one end of the scale (and regardless of their original purpose) these technologies have been employed by various regimes to supress democracy, target activists, instil fear, and prevent social change. While at the other end they have been used to try and prevent the sharing of files on the internet, or even the accessing of websites which provide links to where information is stored; thereby upholding the law while protecting commercial interests and/or enforcing moral standards. This causes problems when search blockers deliberately or inadvertently prevent access to information which is not illegal.

Challenges: To create technologies:

- Which can achieve legitimate security, public disorder, and crime prevention goals without inadvertently providing tools of oppression for brutal governments.
- Which citizens trust to the extent that they will not be fearful of, or dissuaded from, exercising their freedom of expression.
- Which do not block access to lawful information, even if this information is controversial.

Step 1. Identify specific scenarios where it will be ***lawful*** to use your proposed security technology when it is having the effect of infringing someone's freedom of expression?_____

If you cannot identify any such circumstances then chances are it will be unlawful to use your proposed technology.

Step 2. Identify scenarios where you think it will be ***unlawful*** to use your proposed security technology when it is infringing someone's freedom of expression?

What protections could you possibly build into your security technology to either prevent or minimise the likelihood of such unlawful use occurring?

If your technology is open to unlawful use and there are either no technical protections available, or there are but you are not going to build-in these protections, justify why your technology should still be built.

Step 3. If your were personally being monitored by (or subjected to) your proposed security technology would you be as likely to, or worried about, engaging in the following legal activities (which for these purposes you wish to engage in): attending a meeting of the BNP?_____; organising or participating in a protest rally?_____; writing a blog criticising the government or your employer?_____; hosting a Wikileaks style website?_____; publishing a newsletter promoting controversial views (i.e., anti-homosexuality views based on religious doctrine, criticisms of British soldiers, assisted suicide, extreme religious views on the role/rights of women, anti-abortion, pro-abortion, promotion of the death penalty, etc)?_____. If you answered 'yes' to any of these your technology is having a chilling effect. How can you minimise this effect through your design?

Note on Step 3: Freedom of expression does not mean 'the freedom of others to only express uncontroversial views that you agree with'. It means the right to express views which others may find offensive, disgusting, even dangerous; views which go against everything either other individuals or the majority stand for and have fought for.

Qn.2e: Could the security technology unintentionally affect your right to a fair trial or decision?

When I talk about *trials* and *decisions* here I am not restricting this to courtrooms and formal criminal trials. I am referring to the processes by which a decision is reached which materially affects the target of a security technology. This may occur at the moment the technology is applied to the target (e.g. an airport body scanner highlighting something of interest on the body of a passenger), or it can have a delayed effect which might only arise after some action by the target (e.g. a person being included on a no-fly-list as the result of a data-mining process - a decision they may only become aware of the next time they attempt to fly and are refused permission to do so).

It goes without saying that no scrupulous security technology developer (such as yourself) would ever set out to deliberately design and built security technologies with either the primary or auxiliary purpose of undermining a citizen's right to a fair trial or decision. Indeed if such a technology was created no democratic government which respected human rights would deploy such a technology, and if it ever was deployed it would not be afforded legitimacy by the public and would certainly damage the reputation of the developer.

So why should you be concerned about an outcome you are not intending to create? The answer is 'because this outcome can occur as an unintentional side-effect of how a security technology achieves its intended goals'. Technologies where the internal protocols do not begin with a presumption of innocence for the target, where the decision making processes are opaque and/or not set out, where the target is not afforded the opportunity to defend themselves and/or not presented with the evidence against them, or where no prior reasonable suspicion existed before they were subjected to the technology, can all affect a person receiving a fair decision.

If this Section is applicable you should look at Section 3c for the interrelated topic of citizens being able to determine the legality of how a security technology operates.

Step 1. Consider the scenario below before answering the questions which follow.

Your proposed security technology has been completed and deployed. An individual (Louise) has been subjected to the security technology. As a direct result of this she was identified as a threat and relevant sanction(s) have been applied to her (the form of these depends upon the nature of your security technology but they may include; loss of current employment, refusal of future employment, detention in prison, home detention, restriction and/or monitoring of movements, prevented from flying or

travelling, prevented from contacting certain people, internet restrictions, etc.). Louise seeks to challenge the decision against her. Answer the following questions basing your answers on the how you envision your proposed security technology operating.

Why was Louise targeted by your security technology? Was it (a) because she must have sought to enter some area or access some service where everybody is targeted (like an airport)?; (b) because she must have performed some illegal or suspicious action which triggered an alarm or attracted the attention of the technology or its operator?; and/or (c) could she have been selected as the result of a random or mass search for which there was no prior action or evidence against Louise which could have equated to prior reasonable suspicion?_____. If your answer includes (c) then your technology may be conducting *fishing expeditions* (i.e. investigating an individual or mining their data in an effort to uncover evidence of any illegal or suspect activities by that person when you did not possess any prior evidence of wrongdoing constituting reasonable suspicion before your investigation/mining commenced – an action which undermines the presumption of innocence and has led to social and judicial resistance in the past). Describe here what design changes can you make to prevent your technology being used to conduct fishing expeditions? or if you are not going to introduce these changes describe why your technology should still be built._____

_____

_____

Step 2. The following questions are to determine how easy/possible it is for Louise be defend herself against, or challenge the validity of, any sanctions against her.

- Will Louise be shown the evidence against her?_____.
- Will this evidence be presented in such a way that it explains clearly and exactly why any decisions against her were taken at all?_____.
- Will Louise be able to challenge any sanctions/restrictions that are going to be applied to her at the immediate moment that decision is taken?_____.
- Will the operator of the security technology possess the authority and physical ability to override any decision made by the security technology upon hearing Louise's arguments and weighing them accordingly?_____.

For every 'no' answer you make to these questions, if it is later shown that Louise was not guilty of any offence and should not have been subject to any sanctions as determined by your security technology, then it becomes increasingly likely your technology will lead to resentment and resistance by the public. How can you modify its design so as to minimise this risk?_____

_____

_____

_____

Qn.2f: Do the security benefits outweigh the losses to other human rights?

We as citizens have a right to safety and security. Indeed it is often stated that the provision of security is the single most important role of the state, and that security is the paramount right upon which all other rights are dependent; i.e., a secure environment is a prerequisite for the enjoyment of other rights. Nevertheless, as one of many rights, security does not automatically trump these others when they come into competition. Indeed when security is afforded priority over other rights this raises the possibility of social resistance towards the security measures involved. Should the resulting security benefits be outweighed by losses to other rights then the risk of social resistance is not only heightened, it is also afforded legitimacy. It is therefore important that the developers of security technologies ensure that the gains in security arising from their proposed technology outweigh any concomitant losses experienced by other rights.

There is no universal formula for determining whether security benefits outweigh the losses to other rights. Two of the greatest challenges you face here are that:

1. Unlike some of the other *balancing* scenarios addressed throughout this tool at no stage will you be comparing like-for-like values; security is different to privacy which are both different to association, which are all different to religion, etc.
2. Because all the rights are inherently different you will need to convert them into common denominators if you are to measure the loss/gain effects. Additionally citizens may assign different utility values to changes in the quantum of a right depending on the amount of that right they both began with and will be left with. This process *will be subjective* and your methods and results will be open to scrutiny.

Finally just because the security gains from a security technology outweigh the losses to other rights this does not necessarily mean the technology is acceptable. Certain rights may be considered off-limits to citizens regardless of the security gains to be had. If there is only a small amount of a right left for a citizen to enjoy then they may be less likely to accept security technologies which would eat into this right. These are all factors which need to be taken into account when arriving at your final assessments.

Step 1. Identify the rights (other than security) that will be curtailed by the introduction of your proposed security technology as well as the rate (nature and scope) of these curtailments?

Step 2. What are the anticipated security benefits and their rates?

Step 3. Are the 'curtailment rates' (Step 1) and 'security benefits' (Step 2) common values which can be compared?_____ If 'yes' compare these two values. If 'no' first convert them into a common denominator which can be compared before going ahead and comparing these two values here.

Note on Step 3: Be aware that citizens may vary the value they place on a right differently depending on how much of that right they started with and how much they will be left with. Also certain rights may be considered *off-limits* such that no security gain will be acceptable if it requires this right be lost or even infringed.

Step 4. Do the security benefits outweigh the curtailments? If 'yes' what concerns remain regarding this calculation? If 'no' how can development and deployment of this technology be justified

Qn.2g: Could the security technology be used for torture or inhuman or degrading treatment of the targets?

Article 3 European Convention on Human Rights states in simple terms that no one shall be subjected to torture or to inhuman or degrading treatment. This is the only right within the Convention that is unqualified. In other words this is the only right that is absolute; there are no opt-outs, circumstances, or competing reasons available to governments allowing them to lawfully infringe upon this right.

There is therefore little point in security developers expending resources (time, money, effort, etc.) designing torture devices as commercial products for few governments will be able to use them; these devices will probably be subject to trade restrictions anyway; and in any event once the public find out it will almost certainly mark the end of that developer with the possibility of criminal sanctions for the individuals involved.

This is also an area where developers run huge risks should they seek to exploit dubious grey areas in definitions. For example; waterboarding is *torture*! It is not merely an *enhanced interrogation technique*. Designing a better waterboarding device to sell to the United States and hiding behind such linguistic justifications will not be acceptable to the public, will most probably mean the end of your company, as well as opening the door to criminal sanctions.

Given that you are not intending to develop torture devices the major concern in this area is the dual-use of security technologies which were designed for other purposes but are subsequently used as instruments of torture. The most likely candidates here are less lethal weapons (i.e. teargas, truncheons, water cannons, Tasers, baton rounds, etc.). Of these candidates Tasers are probably the most concerning as they are widely issued, have a history of misuse in police stations away from the gaze of the public, leave no lasting marks, and use pain to both subdue citizens and force them to comply. These are devices which are used as torture instruments throughout the world today.

Challenge: To design security technologies where; (a) it is either impossible for them to be used as instruments of torture or inhuman treatment in addition to their *proper* intended purpose, and/or (b) where misuse of these technologies is automatically recorded so as to deter the user from such actions for fear of the repercussions.

Step 1. Are you designing your proposed security technology with the intention that it can be used as an instrument of torture, either as a primary or secondary function?_____ If 'yes' how can you justify producing this technology given the massive risks (social, reputational, legal, ethical, financial, etc.) this entails?

If 'no' continue on to the remaining Steps in this Section.

Step 2. Consider your proposed security technology. Can you conceive of any possible methods/scenarios where your technology could be used as an instrument of torture?_____ If 'yes' list these out here, and while doing so make sure to note for each method whether this capacity for torture is '*technology based*' or '*end-user based*' in relation to the anticipated final product.

Note on Step 2: What I mean by 'technology or end-user based' in relation to the anticipated final product' is whether the technology only becomes a torture instrument because of the actions of the *end-user* (such as a police baton), or whether this is also the result of the some capability of the final product (such as a Taser with an inbuilt *Drive Stun* mode) which for our purposes is *technology based* even though it also requires actions by the end-user.

Step 3. For those methods/scenarios identified in Step 2 where your proposed technology possesses an *end-user* capacity to inflict torture, what design changes can be made to; prevent this occurring, reduce the incidents of occurrence, and/or reduce the effects if an end-user attempts to use your technology to inflict torture?

Step 4. For those methods/scenarios identified in Step 2 where your proposed technology possesses an *technology-based* capacity to inflict torture, what design changes can be made to; prevent this occurring, reduce the incidents of occurrence, and/or reduce the effects if an end-user attempts to use your technology to inflict torture?

Note on Steps 3 & 4: Such protections might entail: methods for monitoring and recording the use of the technology, restrictions on how many times and how often the technology can be used, designing the technology to physically break if a certain amount of force is applied by the end-user, reducing the power/effects of the technology, etc.

section 3:

questions of legality

Qn.3a: Is the use, sale, or possession of similar security technologies restricted or restricted?

It is not uncommon or controversial for security technologies to be subjected to restrictions governing who can legally possess and use these technologies. Similarly the sale of security technologies is often restricted and regulated. This reflects, amongst other things; the dual-use potential of many security technologies whereby if used inappropriately they can often cause harm as readily as they can provide a security benefit, the health and safety risks inherent to certain security technologies, and the need for special training. In certain instances the possession, use, or sale of a security technology is banned completely.

Security technology restrictions also reflect the views of society. Pressure placed on governments (both local and national) and regulators by both the general public and specific activist groups has led to the restriction and/or banning of certain security technologies in the recent past.

- The UK's National Identity Register and its associated National Identity Card were abandoned following the ultimate display of societal power – the election of opposition political parties with manifestos to destroy these programmes.
- The use of Mosquitos has been banned in certain local authorities following complaints to and from local councillors.
- The sale of the fake handheld explosive detector *ADE-651* to Iraq and Afghanistan was banned following media investigations and the ensuing public response.

It is in the interests of both those designing security technologies and the wider society that security technologies with the potential to provide real security benefits can be developed and deployed in a manner which society finds acceptable.

Challenge: To produce a security technology which will not be banned and will only be subject to reasonable and proportionate restrictions. To be achieved by identifying those design characteristics of similar banned/restricted security technologies which have justified these bans/restrictions and then using this information to inform the design of any new security technology.

Step 1. Are there similar technologies to your proposed technology, both by design and effect, already on the market?_____

Step 2. What restrictions are placed on the sale, possession and use of these technologies?_____

Step 3. What are the design characteristics of these technologies that are related to or influence these restrictions, and how do they have this effect?

Step 4. Consider the proposed design characteristics of your own security technology. Do any of these overlap with the specific design characteristics of existing technologies identified in step 3?_____ If 'yes' move on to steps 5 & 6.

Step 5. List off the overlapping design characteristics identified in step 4.

Note on Step 5: Ensure that throughout the design process this list remains highlighted to the designers. It sets out opportunities to produce an innovative security technology with the potential to:

- Be acceptable to the public
- Maximise the uptake of this technology
- Minimise both the likelihood and level of restrictions on its use and sales

Step 6. How can you address these potentially controversial design characteristics in your proposed security technology?

Qn.3b: Are any design or operational elements likely to provoke legal challenge?

This section is not about determining whether or not a technology *is* or *isn't* legal. Rather it is about attempting to assess the likelihood somebody will challenge that technology in court.

Predicting with accuracy whether a security technology possesses elements likely to provoke some future legal challenge can be a difficult task, unless that technology is blatantly illegal and/or carrying out its functions in an illegal manner at which point it is almost certain to face legal challenges by its targets. In any event, given the possible downsides of having a security technology declared illegal by a court, it is worth investing time in this endeavour.

Challenges to the legality of a security technology can take considerable time to resolve. Court cases attract publicity and, depending on the outcome, can force both developers and end-users to make changes to the design of a technology and/or how and when it may be used, assuming the technology remains legal to all. Even if the technology wins, negative publicity preceding the trial may undermine the technology in the eyes of the public. Indeed cases need never go to trial for there to be a detrimental impact; by raising specific concerns of legality which will remain moot until there is a court ruling questions will remain over the standing of the technology.

For example:

- The use of backscatter whole body scanners on children in airports was temporarily halted in response to claims the images produced were in violation of the Protection of Children Act 1978. By the time these scanners were removed from UK airports no case had been brought on this point.

We have adopted two approaches for achieving the aims of this section:

- Encouraging the developers to step back and consider the proposed technology from the position of the different people subjected to it
- Examine the proposed technology in the light of historical examples where elements of a security technology have raised the threat of legal challenge.

Challenge: To identify those statutes that may apply to the proposed technology, and to anticipate how people who are being subjected to the technology will react. Then using this information to either reject changes, or to modify the proposed design and/or build in protections.

Step 1. Sketch out your proposed technology; focussing on who will use it, where it will be used, what it will actually do, and how it will do it.

Step 2. Use the information from step 1 to identify those statutes which may apply to this technology. For each statute formulate a plan of action: this may be to reject the statute, modify the design, reject elements of the security technology, lobby for law changes, etc. Justify your decision making process for each of these statutes.

Notes on Step 2: Without prior experience in the regulatory field surrounding the proposed security technology it is advisable here to seek professional legal advice to identify the specific laws which may apply.

Step 3. Consider possible responses to the proposed technology from the perspectives of members of the specific groups listed below. Assume they have been subjected to the technology and that it was used correctly and according to any relevant codes of practice by appropriately trained staff:

Physically disabled:

Mentally disabled:

Parents:

Children:

Pregnant women:

Elderly:

Ethnic minorities:

Religious groups:

Step 4. Repeat the process from step 3 but now consider the response if the technology is being misused, abused, used without proper protections, used by staff who are not properly trained or are acting inappropriately. How will they respond under these circumstances?

Notes on Steps 3 & 4: The aim here is for the developer to mentally step back from their design. You may believe that what they are developing is completely justified because you can see the security benefits. However, it is how it interacts with society and social groups which will ultimately determine the acceptability and uptake of your technology.

Try and place yourself in the shoes of somebody from each groups listed above. How would you feel and react if you were the subject of this technology, both when it is used properly but also if it is used inappropriately and/or deliberately misused by the end-user. Make sure you include extra groups to the list above if you can foresee the potential for specific issues for specific groups arising from your proposed technology.

Qn.3c: Could an individual determine whether the security technology targeting them is being operated legally?

The operation of security technologies is often accompanied by a level of secrecy so as to hopefully enhance their effectiveness. For example profiling, data-mining, and data-matching technologies designed to increase the detection rates of terrorists and other criminals based on certain activities and/or attributes are only effective if the terrorist fits the previously created profile. Proponents of secrecy argue that if this profile is openly known then it becomes easier for terrorists to avoid the actions which will enhance their likelihood of being detected and/or recruit certain types of individuals who are less likely to trigger alarms.

This secrecy is often not problematic providing society can have trust in how these technologies are formed and operate. This trust requires these technologies operate within the confines of the law. Security technologies which discriminate by religion, nationality, ethnicity, disability, age, sex, etcetera, are likely to run afoul of the various anti-discrimination Acts as well as the Human Rights Act 1998 regardless of their security purpose. The public's trust will also be called into question by administrative decisions taken exclusively by a security technology without any human oversight – especially when that decision is patently wrong.

Finally there is the need for the decision making process to be open to scrutiny by those affected by it and who disagree with the decision taken. Hiding behind commercial secrecy or national security to prevent the open disclosure of this process, especially in court, will naturally call into question the integrity of the security technology.

Challenge: To develop security technologies which may be subject to secrecy requirements, take independent decisions, and/or employ complex decision making algorithms, but which can also provide those affected by it the information they need to determine whether or not that technology was operating legally. This includes providing a clear rationale for why they were targeted as well as the justification behind the decisions taken.

Step 1. Will information about the operation of the proposed security technology (and where relevant the decision making processes including decision algorithms) be made available to the public?_____ If 'no' what is the justification for adopting this approach?_____

_____

_____

_____

If 'yes' what information will be made available?_____

Step 2. How is the information presented to the individual from step 1 sufficient to allow that person to determine whether or not the security technology was operating in a legal manner when they were subjected to it? Alternatively, if it is not sufficient how do you justify the level of detail provided?

Step 3. Is the information provided regarding any decision making processes understandable and opaque to an individual who doesn't possess a technical background?

Step 4. How are administrative decisions made by the security technology communicated to the target? When does this occur? And are these decisions challengeable by the target before they can have an adverse effect on that individual?

Step 5. How can the proposed security technology be designed so as to prove that it was operating in a lawful manner at the point in time when it targeted an individual?

Qn.3d: Is illegal use of the security technology prevented or preventable by design safeguards?

There is a multitude of security technologies built to protect individuals, property, and/or the state. Successful technologies can enhance our collective security by detecting, preventing, and deterring crimes, and catching offenders. Despite these benefits possibly all security technologies could be classified as *dual-use* technologies; possessing both the capacity to increase security (as intended by design) as well as the capacity to assist in the carrying out of crimes. Tasers used to immobilise an attacker have also been used as instruments of torture, CCTV cameras which record crimes for future prosecutions but also have been used to spy on residents, ID cards for proving a person's identity can also perpetrate identity fraud, and so on. It is an inconvenient fact that some end-users in the security field (such as police and government officials) abuse their positions, access, and technologies to commit crimes.

Security technologies which are widely viewed as open to abuse and/or are indeed used as tools of abuse will not enjoy the support of citizens. While of course not all security technologies can be designed to prevent their illegal use, for those where this is possible the building-in of such protection could only be beneficial.

Challenge: To build protections into the proposed security technology that will prevent and/or deter the misuse and illegal use of that technology. This could be by protections which prevent the misuse at the time of use (*present misuse*), or by protections which record instances of misuse of that technology, thus providing evidence for future action against the user (*future misuse*).

Step 1. Is it possible to build protections into the technology to prevent the present or future misuse of that technology?_____ If 'yes' continue on to step 2. If 'no' outline why this is impossible_____

_____

_____

_____

_____

_____

_____

Step 2. Describe what measures could be designed into the technology to prevent that technology being used illegally (i.e. to prevent *present misuse*)?_____

_____

_____

Step 3. How could use of the technology be monitored so as to prevent misuse and/or prosecute users for past misuse (i.e. *present* and *future misuse*)?

Notes on Steps 2 & 3: Possible approaches here may include measures:

- To prevent the technology from physically being used by an authorised person(s).
- Whereby the technology records on a log when it was used, by whom, and how.
- The in-building of privacy enhancing technology into surveillance systems to restrict what can be viewed.
- An alarm system build into the technology to warn others in real time that; (a) a technology is being used by an end-user, and/or (b) that there is evidence to suggest it is being used illegally or inappropriately.

Step 4. Are there specific obligations on the user to take positive actions to prevent the technology being misused or used without authorisation? If 'yes' what protections could be built into this technology to achieve these aims?

Qn.3e: Does the security technology respect basic legal principles within society?

Basic legal principles have formed as both our society and our legal systems have developed. Collectively these principles may be termed the *rule of law*. While it is tempting to skip past such ideals as too hard to define and/or too uncertain and subjective in nature, this is inadvisable. Concrete requirements of operation and fairness can be identified through an examination of the rule of law, and the application of these basic requirements will only strengthen a security technology, while ignoring them will potentially undermine social acceptability of that technology.

Relevant elements of the rule of law are listed here (note that a number of these are examined in greater detail in other sections so will not be addressed in this section). We have modified these so as to apply them to the designing of security technologies. These requirements and their subsequently modified formats are as follows:

- (*rule of law*) The law must be accessible, intelligible, clear, and predictable; (*design principle*) The legal basis for the security technology should be clear.
- (*rule of law*) Legal rights and liabilities should be determined by application of the law, not the exercise of discretion; (*design principle*) Decision making processes of a security technology should not be arbitrary. It should be clear with a defined scope.
- (*rule of law*) Adjudicative procedures provided by the state should be fair. (*design principle*) Decision-making procedures of a security technology should afford those listed as targets the opportunity to challenge this claim.
- (*rule of law*) Public officers should exercise their powers fairly, for the purpose for which they were given, and without exceeding these powers. (*design principle*) Security technologies should be used for approved purposes and must not exceed the powers assigned to them.
- (*rule of law*) The law should apply equally to all (see Question 3f for an examination of discrimination).
- (*rule of law*) The law must protect human rights (see Section 2 for examinations of human rights)

Step 1. Is there an existing legal and/or regulatory system governing the use of the proposed security technology?_____ If 'no' see the following Note on Step 1. If 'yes' does the proposed technology meet these requirements?_____ If 'no' to this second question what changes need to be made so as to ensure compliance?_____

Note on Step 1: While the absence of a regulatory system governing the design of a security technology may allow a designer more freedom in the short term, this can have legal repercussions in the mid to long term. Security technologies that lack a statutory basis cannot rely upon such a basis if challenged in court, and citizens may be more likely to engage in legal challenges of this technology given this fact.

Step 2. How can the design of the security technology prevent arbitrary decision making by either the technology itself or the end-user?

Note on Step 2: Measures may include either automatically-recording, or forcing the user to record both when the technology is used (including the justification for use) as well as when the technology is not used (as well as the justification for not using it).

However, there is a difference between arbitrary decision making and the use of human discretion/intuition. When the decision of an automated security technology should be overridden by a human operator because common sense dictates that it should be, then this must be allowed though details of such actions should be logged.

Step 3. How does the security technology provide opportunities for individuals to challenge the output/decisions of this technology?

Can they challenge this decision/output immediately after it is produced so as to minimise the effects of an incorrect decision? If 'no' how long will the effects of an incorrect decision be applied to the target, and how can the design be modified to reduce this time as far as possible?

Step 4. What are the approved security purposes for the proposed security technology and how can the design of the technology ensure that the technology is not used for unapproved purposes?

Qn.3f: Can the operation of the security technology entail discrimination?

Discrimination against individuals based on some trait, belief, or inherent characteristic is rightly illegal within many countries. This is unsurprising given its direct relationship with some of history's darkest moments including slavery and the holocaust. Discrimination encapsulates racism, sexism, ageism, religious intolerance, and disability discrimination to name some of its forms. Within the United Kingdom there are various pieces of legislation banning discrimination including, but not limited to, the Race Relations Act 1976, Sex Discrimination Act 1975, Disability Discrimination Act 1995, and the Human Rights Act 1998. The fact there are so many Acts should leave designers of security technologies in no doubt as to the unacceptability of this outcome.

It is also clear that *security* does not trump the prohibition on discrimination as the example below from aviation security illustrates:

- The Interim Code of Practice for the Acceptable Use of Advanced Imaging Technology (Body Scanners) in an Aviation Security Environment explicitly states, "*Passengers must not be selected on the basis of personal characteristics (i.e. on a basis that may constitute discrimination such as gender, age, race or ethnic origin)*".

The onus is on designers of security technologies to ensure their technologies do not directly discriminate against different groups of citizens. Additionally, given the likelihood of a social backlash against technologies which facilitate discrimination, designers should also be seeking to minimise the possibility end-users could use their technologies in a discriminatory manner.

Challenge: To design technologies which do not directly discriminate while also minimising the probability/instances of end-users discriminating with the technology.

Note on *Profiling*: Profiling refers to both the process of building profiles (i.e. discovering correlations between data so as to create a representation of a group) and the process of trying to match individuals to a previously created profile. The goal is to identify individuals with the highest probability of being a criminal before attacks occur. The dangers here are that; (1) the data used to form the original profile is discriminatory; (2) the profile is discriminatory in composition; (3) those trying to match individual citizens to a profile do so in a discriminatory manner. While profiling does have security benefits it always walks the tightrope of constituting discrimination which will undoubtedly result in social resistance to this technology and a lack of legitimacy.

Step 1. Begin by describing both how your technology acts upon an individual as well as how the operation of this technology will be experienced by the typical citizen-target. Does this technology discriminate against the typical citizen, and if so how?_____

Note on Step 1: By 'how your technology acts upon an individual' we mean everything your technology does in relation to a target when it is operated. For example; for a data-mining programme this will involve what information is searched, what criteria were used when searching, what is chosen and discarded, and how it is collated and presented; for a tear-gas canister it is how the canister is deployed and how the gas effects an individual, how long it last, etc.

Step 2. Using your descriptions of how the technology acts from Step 1, replace the 'typical citizen-target' with an individual from each of the following eight groups; again asking whether the technology discriminates against any group, and if so how?

children	male/female/transgender
different religious groups	foreign national
elderly	physically/mentally disabled
different ethnic groups	different skin colours

Step 3. If discrimination is identified in Steps 1 or 2, how can the design of your potential security technology be modified so as to minimise and ultimately prevent this discrimination? Alternately, if you choose not address this discrimination justify your decision here.

Step 4. How can you build protections into your proposed security technology to minimise and ultimately prevent end-users discriminating when they use the technology?

Note on Step 4: Measures here might include: logs recording when the technology was used, by whom, and for what purpose; a capability whereby the target is provided information about why they were targeted as well as the history of use of the operator; a real-time oversight function; the public disclosure of information regarding how the technology is being used, etc.

Qn.3g: Will there be design protections within the security technology to ensure compliance with all data protection principles?

The advent of digital databases has facilitated the collection, storage, and processing of personal data on a scale never previously achieved in human history. These systems have the power to contribute to economic and social expansion, facilitate trade, and improve the well-being of both individuals and society.

However they also possess the potential to threaten the individual. They can undermine the fundamental rights and freedoms of individuals, most notably their right to privacy. Additionally, databases of personal information are targets for attackers who can use this information to commit identity fraud, theft, and other crimes.

It is against this conflicting backdrop of benefits and threats that has seen data protection principles enshrined in law through European directives (such as *Directive 95/46/EC*) and associated national legislation (such the UK's Data Protection Act 1998). These provide principles governing the collection, storage, and processing of data; all designed to protect an individual's personal data and to rein in the otherwise unrestricted collection and processing of such information.

It must be noted that these principles and restrictions are often presented with the attached caveat that the scope of some of these obligations may themselves be restricted by governments when the collection and processing of personal data is *necessary* to safeguard national security, defence, safety, and/or the prevention, investigation, detection, and prosecution of criminal offences.

However caution is advised here, for these caveats do not afford developers or end-users *carte blanch* to ignore data protection principles. On the contrary, measures adopted which are not *necessary* will still be in breach of these principles. Similarly, regardless of legality if a proposed security technology appears excessive or disproportionate in its collection and associated processing of data, the real possibility arises it will lack public support and/or evoke active public resistance and thereby lack societal legitimacy.

Challenge: To create a security technology that possesses design features allowing it to:

- Comply with all data protection principles during normal operation;
- Identify and facilitate the wider collection and processing of data when such actions are legitimately necessary to safeguard specific goals (i.e. national security, public security, defence, preventing crimes, etc.);
- Provide the public with the relevant information so as to assure them that the technology is both operating and being operated in a legitimate, acceptable manner.

Step 1. Answer the following series of 'yes/no' questions:

- Has consent been given to the collection or processing of data?_____
- Are individuals informed when their personal data is processed?_____
- Is personal data anonymised?_____
- Does the ability exist to block access or delete unlawfully processed data?_____
- Does the ability exist to rectify data that is wrong or incomplete?_____ and can this happen immediately upon discovery of the errors?_____
- Are *specific purposes* specified each time data is processed?_____ and are these purposes recorded for future examination?_____
- Are *legitimate purposes* specified each time data is processed?_____ and are these purposes recorded for future examination?_____

For each 'no' answer above, identify and write down design measures and which could be implemented or will need to be addressed to achieve what is set out in the relevant question(s). Alternatively, if you have decided not to address any of the 'no' answered questions set out your justifications for doing so._____

_____

_____

Step 2. To ensure *proportionality* of the proposed security technology, and prevent the excessive collection of information; firstly set out what security problem(s) your technology seeks to address. Then for each identified problem set out what you consider to be the *minimum-necessary* amount of data on an individual which would need to be collected to address that security problem._____

_____

_____

Consider now your proposed security technology. Does its data collection exceed these *minimum-necessary* limits?_____ If 'yes' how can this be reduced through the design?_____

_____

What design protections can be put into place to ensure future data collection does not exceed the *minimum-necessary* data collection limits when the security technology is deployed?_____

_____

Step 4. What measures are built in to protect the data that is held?_____

_____

If specific protection obligations exist do these measures meet them?_____

Step 5. How is use of the security technology monitored to prevent illegal use by both insiders and external attackers?_____

_____

_____

section 4:

financial cost of the ST

Qn.4a: Do the financial costs outweigh the security benefits?

The development and implementation of security technologies where the financial costs outweigh the security benefits (either in reality or by appearance) have caused controversy in the past. Especially when systems are paid for by the public through taxation, any failure of the technology to deliver commensurate security benefits is at risk of attracting negative social and media attention.

There is no universal formula for determining whether security benefits outweigh total financial costs. While in some cases a simple balancing of equivalent values (money saved versus costs of the security technology) may be possible, any calculation in this area will depend upon the characteristics of the benefits produced. Lives saved, injuries prevented, industries protected, jobs maintained, fear reduced, and business continuity maintained, etc., will all need to be considered and converted into common values so as to be weighed against each other. However when the values are not ostensibly equivalent (i.e. money versus lives saved; money versus reduction in fear of citizens) then the user of this tool will need to produce an explicit formula for converting these variables to a common denominator.

When presenting the financial cost of a security technology as a justification for the introduction of that technology, do not simply use the *per unit cost*. Factor in all one-off and on-going costs, including maintenance, staffing, installation, training, etcetera; which I refer to as the *true purchase cost*. This information will almost certainly come out into the public domain, especially if the technology is potentially controversial for other factors. Any slow drip-feeding of information (whereby the price of the technology constantly rises) will only increase suspicion as to the actual cost. Undervaluing the *true purchase cost* will not change how much it actually is, and it is this cost which will ultimately need to be acceptable to the public.

Finally, just because the security benefits of a security technology are outweighed by its financial costs does not mean this technology is necessarily unacceptable or not worth implementing. Factors other than cost may make the technology acceptable, such as providing comfort, protecting jobs and services, etc. These all need to be factored in to any final calculation.

Step 1. Begin by setting down what the *true purchase costs* of the technology will be for the client assuming they wish to purchase enough units so as to effectively use them throughout their organisation (which for a government may be an entire nation). This figure includes all one-off set up costs for installing the units, training/accreditation costs for end-users and on-going running/upgrade costs for 1, 5 and 10 year periods. When in doubt err on the high side of any calculations and make

sure you include the effects of inflation in your calculations.

Notes on Step 1: Do not use lower-end/conservative estimations of costs; given the number of past security technologies which grossly exceeded their original design and instillation budgets; the public will not trust low estimates, and any future upward corrections of the price will merely justify and strengthen this opposition.

Set-up costs include physical building changes necessary to accommodate the technology (e.g., airport terminal changes to accommodate whole body scanners)

Step 2. What are the anticipated *security benefits* and their rates?

Step 3. Are the *true purchase costs* and *security benefits* common values which can be compared?_____ If 'yes' compare these costs and benefits. If 'no' first convert harms and/or benefits into a common value which can be compared before comparing harms and benefits.

Step 4. Do the security benefits outweigh the *true purchase costs*?_____ If 'yes' what concerns remain regarding this calculation? (If 'no' go to Step 5)

Step 5. If 'no' how can the development and deployment of this technology be justified?

Note on Step 5: There can be benefits beyond *security benefits* which may justify the financial costs of developing and deploying a particular security technology. For example, reducing fear, protecting industries and/or jobs, maintaining business continuity, protecting reputations, etcetera.

Qn.4b: Will the cost of this technology be considered too high or excessive?

While governments are generally willing to spend considerable money on security technologies, especially when perceived threat levels are high, it would be wrong to assume the public will automatically be as willing to consent to these large purchases. A number of factors can influence the likelihood such financial outlays will meet public resistance:

- If the technology is controversial to begin with for reasons other than cost then a high price-tag will provide a convenient justification and focal point from which to oppose the purchase and introduction of that technology; e.g. see the now defunct National Identity Register and associated National Identity Card system.
- If the public can see they are going to have to pay for it directly, either now and in the future through higher prices/fees, as opposed to the costs coming out of central finances, then there may be greater resistance.
- If the public cannot be convinced that a particularly large amount of money should be outlaid (either on this technology or at this particular point in time), resistance may occur regardless of whether the benefits of the technology outweigh the financial costs.

What is obvious here is that not all of these factors are necessarily within the power of the designer to directly control. For example if governments choose to maintain secrecy over the successful contraband detection rates of airport whole body scanners so as not to encourage attacks then this will impact upon the amount of information which can be released to the public when developers are trying to win the public over. And if trust levels in government are already low then the public may be less inclined to simply believe the government when they say a technology is good value for money, effective, and a necessary investment.

Challenge: To design a profitable security technology while at the same time winning over a potentially sceptical public by proving the necessity and effectiveness of this technology, and doing so when they may not have control over the information-dissemination flows.

Step 1. Start by setting down what the purchase costs of the technology will be for the client assuming they wish to purchase enough units so as to effectively use them throughout their organisation (which for a government may be an entire nation), including any one-off set up costs for installing the units, training/accreditation costs for end-users, and the on-going running/upgrade costs for 1, 5 and 10 year periods. When in doubt err to the high side of any calculations and make sure you include the effects of inflation in your calculations.

Step 2. Take the final figure produced from Step 1 and look at it, not as the designer of security technologies, but as a sceptical citizen who has no vested interest in this technology being successful. Ask yourself the following questions:

- How will this amount appear to citizens given the current environment?
- Looking at it subjectively, does it appear large or excessive?
- How easy would it be for you personally to organise a campaign against this technology if you based it upon the costs involved?
- What else could be purchased for this money? (e.g., how many teachers, nurses, fire-fighters, etc., could be employed with this money) So is it a good investment?

Note on Step 2: For this Step to work you really need to try and place yourself in the shoes of the sceptical citizen. This requires you taking off your technical hat and forgetting about any personal gains you may make from this technology. Be the protagonist. Think of this technology as an invention by your main rival; how easily could you use the cost of this technology to lobby public resistance to its introduction?

Step 3. How will this technology be paid for? (e.g., taxes, user pays, surcharges, etc.), and how could the answer to this question affect citizen's acceptance of this technology?

Note on Step 3: Even if the costs per citizen/user are low, be careful assuming this fact will not lead to resistance, especially if the area involved is already subject to many fees and charges, or alternately if the area involved was traditionally free from charges.

Step 4. Will you be able to freely release relevant information on the capabilities/detection-rates/benefits, etc., of the proposed security technology into the public sphere? If 'no' why not, and what can you release?

Qn.4c: Will the financial figures released into the public domain be trusted?

This question follows on from Qn.4b. It asks whether or not the figures used to justify the security technology will be accepted as both accurate and realistic by the public, and hence be trusted by them.

The importance of this question is based on the assumption that for every security technology there is a financial cost threshold above which the majority of the public will no longer support the introduction of that technology. Of course unless the public are specifically asked beforehand (something which never happens) this value will never be known with certainty before the technology is proposed to the public (or indeed until the threshold is crossed).

When releasing figures on a security technology three decisions need to be made. Firstly whether or not to release any information at all? There are valid reasons for not doing so. This information may well be commercially sensitive, and/or the designer may be constrained by contracts or laws forbidding them from releasing such information.

Secondly, assuming information is to be released, exactly what figures will be released? This opens up many secondary questions: will they only reflect the price per unit?; will they include on-going running costs and/or initial set-up costs?; who will be releasing the figures?; at what stage in the design process will they be released?, etc.

Thirdly, how (if at all) will these figures be verified before/after release? Will they simply be announced by the developing company or the client without further actions, or will they be independently audited so as to give extra credibility to these figures?

These are all important questions. Security technologies in the past have suffered from the inaccurate release of cost information. When initial estimates are unrealistically or naïvely low, they will be easy targets for external experts to debunk. When corrections are constantly higher than previous figures, all those in the process (as well as the technology itself) risk losing credibility.

Not releasing cost information, or releasing inaccurate information, is always a risky tactical ploy and one which is inherently illogical. Firstly the information will almost always leak anyway. Secondly there will be a maximum price that society will accept for a security technology. If the estimated price for a proposed security technology, or the actual price of a completed technology, exceeds this threshold then the technology will not survive in the long-term regardless of the initial outlay; failing to release accurate information merely postpones the inevitable.

Step 1. Do you know the *true final cost* of your security technology?_____ If 'yes' set it down explaining exactly what this figure refers to and entails (i.e., unit price, maintenance costs, set-up costs, etc.). If 'no' set down why this is the case and whether/when you expect to have this information in the future._____

Step 2. Will you have the authority to publically release cost information on your proposed security technology?_____ If 'no' who holds the power to do this, and why does this power not reside with you?_____

Step 3. What costing information will be released, and what will these figures refer to? (i.e., unit price, maintenance costs, set-up costs, etc.)_____

Step 4. How accurate are the figures to be released? If they are inaccurate; what is the cause of this inaccuracy?, will this be explained with the release of the figures?, and why therefore are these inaccurate figures being released at all?_____

Step 5. Will these figures be audited/checked/assessed before being released?_____ If 'yes' by whom, and will this information be released? If 'no' why was this decision taken?_____

Step 6. If you are not releasing pricing information (whether by deliberate choice or you are obligated not to) what is your strategy to respond to the public when this information is leaked causing a negative public response?_____

section 5:

public & end-user acceptability

Qn.5a: Can all aspects of the security technology's design be justified?

Security technologies derive their legitimacy from the public, such that if a security technology is to not acceptable to the public then it will not survive over the mid- to long-term. Similarly if specific aspects of a security technology are not acceptable to the public then unless these issues are addressed through design changes and/or operational safeguards then again the technology will lack public legitimacy.

Given the crucial importance of public legitimacy, when releasing and promoting a security technology within the public sphere it is not enough to simply design a security technology that works (i.e., that provides net security benefits) and then rely upon this fact alone as justification for the introduction of the technology. You as the designer also need to justify your design choices to a potentially sceptical, even hostile, public audience.

Finding a balanced, justifiable design can be very tricky for developers of security technologies. On the one hand there is the temptation to create a technology with the greatest level of functionality and features; something which can do everything and anything and thereby offer the greatest potential security benefits. And yet on the other hand if the features which produce the security benefits are considered excessive in scope and operation then justifying this technology may prove impossible.

Challenge: Being able to predetermine which design elements, outputs, functions of a proposed security technology are most likely to elicit negative social responses so that you are prepared in advance to justify your security technology. Examples of where developers have failed to achieve this in the past include:

- Failing to justify exceeding current international standards within specific security fields such as aviation security;
- Failing to convince the public that the benefits of a security technology outweigh the burdens this technology will bring;
- Failing to justify the functions and features of a technology, especially if they are considered excessive by the public.

Challenge: Maintaining the initiative within the social interactions by releasing information justifying both why you designed the security technology in the way you did and why your invention can do what it does, *before* those design choices and design elements have become a source of scepticism and/or controversy.

Step 1. Begin by listing off (a) the features/functions of your proposed technology (i.e., what it can actually do).

Step 2. List off the features/functions of any similar security technologies already established within the environment, noting whether any of these technologies have or are encountering social resistance.

Step 3. Next set out the features of the regulatory environment (if one exists) within which the technology will operate (including minimum standards, legal limits and safeguards, ISO's, etc.).

Step 4. Using the information from Steps 1, 2 & 3 describe how your security technology compares to others (if they exist) and how it fits within both the regulatory environment and the social environment.

Note on Step 4: Focus on:

- Where/how the technology exceeds others, especially where these other technologies have been controversial;
- Where your technology exceeds minimum standards;
- Where it approaches legal limits (such as data collection, etc.).

Step 5. Using the information from Step 4 set out arguments justifying the presence of every feature, function and capability of your proposed security technology.

Step 6. What are the *hard sells* from Step 5 and why? How can you justify including these? and if you cannot then what is forcing you to persist with these elements in your design?

Note on Step 6: By *hard sells* I mean: those arguments which you do not find convincing; areas where you exceed accepted standards; the presence of functionality not directly related to your security goals, redundant functions, duplicated functions, etc.

Qn.5b: Is there a danger of losing public trust in your security technology?

Trust from the different groups that make up society is an essential component of society accepting a security technology. Technologies which lack public trust are consistently questioned by the public, often targeted by public pressure groups.

What makes this fact particularly challenging for developers of security technologies is that trust is easy to lose. There are multitudes of different ways in which the publics' trust of a technology can be undermined, and many of these are by actions of somebody other than the developer. Actions by end-users, governments, and state officials can all undermine public trust in a security technology. Additionally, both the way in which information on the technology is released as well as the composition and reliability of that information constitute a significant component of the trust related issues.

Below I have listed ***9 Commandments for Attaining and Maintaining Trust*** produced through an examination of controversial security technologies. This is by no means a definitive list of all the ways trust in a security technology can be undermined but they should assist developers in avoiding some of the pitfalls of the past when producing and presenting their technologies.

9 Recommendations for Attaining and Maintaining Trust

1. Do not deliberately withhold information on how your technology is being used and how successful/unsuccessful your security technology is from the public.
2. Release experimentation/testing data to prove the safety and efficacy of your security technology. If it is not safe or effective then say so.
3. Do not: lie about your product, spin its capabilities, make claims which are not fulfilled, make promises which are not kept, or exaggerate product's potential.
4. Ensure information released is accurate and released promptly so as to avoid the *drip-feeding* of *truthful* information in response to prior misinformation or the absence of information.
5. Be open from the start about the risks inherent to, and created by, your technology.
6. For technologies which include automated decision making capabilities, make sure the rationale behind decisions reached is clearly spelt out, and that the decision is presented in a form which can be challenged.
7. Within the design, where possible include capabilities to prevent the misuse and abuse of the security technology.
8. Within the design, include a capability which records and/or independently monitors whenever the security technology is used to deter and minimise misuse.
9. Governments which forcibly introduce a security technology must protect citizens from the misuse or abuse of that technology.

Step 1. For each of the *9 Recommendations for Attaining and Maintaining Trust* answer the following questions in turn: Will you be abiding by this commandment? If 'yes' what is your plan for doing so? If 'no' why are you not doing so?

Step 2. Consider your proposed security technology. Can you identify any other reasons or scenarios specific to your technology which may have the potential to undermine public trust in the technology? If 'yes' what are these and what can you do to mitigate this effect?

Step 3. If you are not able to meet all 9 Commandments and/or you have identified other reasons why public trust may be undermined, consider the combined effect of all these outcomes. Given these combined effects, do you foresee a risk that the public will not trust your proposed security technology?

Step 4. What are your overall strategies for obtaining, maintaining, and building-upon the public's trust, regarding the design of your product and in its on-going use?

Note on Step 4: You do not need to produce detailed strategies here. Merely sketch out at a high level the types of actions and activities you are proposing to try and gain the public's trust such as; public engagement exercises, publicity campaigns, the establishment of independent oversight controls, etc.

Qn.5c: Is there a danger of losing end-user trust in your security technology?

End-users are those who actually have to make your technology work in the field. Sometimes this will be from a location physically removed from the security technology itself such as the control room for a CCTV unit. Sometimes both the security technology and end-user remain physically separate from human targets such as a security operator running a data mining programme. But other times both the end-user and security technology will be in physical contact with the human targets, such as Tasers and whole body scanners.

In all these scenarios there is a human operator in the form of the end-user, and they need to trust the security technology if they are going to use it *and* believe its output. The results of technologies which produce too many false positives will be ignored by the operators and eventually not used at all. Technologies which produce too many false negatives will similarly be ignored and will likely be replaced by better technologies/methods in the future, especially in the face of successful attacks which the technology failed to prevent.

Even technologies which purport to be automated (such as RFID card access doors to a building or automated threat detection systems attached to CCTV networks) do not exist and cannot operate without human supervision or interaction, such that if the human end-user does not trust the system it will be ignored or circumvented. For example RFID card access systems can fail to recognise legitimate cards, cannot think independently to allow access to a person who has forgotten their card, suffer from electrical faults, and/or can be tricked by fake cards, therefore the system will require a human supervisor if it is to operate effectively. Similarly, automated threat detection systems cannot physically *check-out* a possible threat; they cannot open an unattended bag and search it, nor question a person who has inadvertently entered an area of an airport where they do not have permission to be. These require human end-users to assess and give effect to the decisions made if the system is to operate effectively.

The point of this discussion is that human end-users are a necessary part of security technologies and without their belief and support in the technology it will be ignored, circumvented, or simply not used. Any lack of support *will* find its way into the public domain via media reports, interviews, anonymous sources, public pressure groups, or by those citizens who are subjected to a security technology and witness first-hand the lack of trust and belief the end-users have in that technology. Thus public support and end-user support cannot be separated.

Step 1. What end-user engagement/advice/opinions/feedback have you collected or are you going to collect before the design of the security technology is locked-in?

Step 2. If you are not engaging end-users what are your reasons for not doing so?

Step 3. Have there been examples of previous security technologies similar to yours which have been criticised by end-users? If 'yes' identify the reasons for this criticism and explain how the design of your proposed security technology will attempt to address these criticisms.

Step 4. Are you going to instigate a system for collecting end-user feedback once the first prototypes of the security technology are produced but before the technology is unveiled?

Step 5. Will the problems of false positives and/or false negatives potentially apply to your proposed security technology? If 'yes' complete Step 6.

Step 6. How many people will be subjected to the proposed security technology? What will be acceptable false positive and false negative rates? How many people do you therefore predict will therefore become false positives and false negatives on any given day? What will be the predicted effect of such numbers on the end-user's trust and belief in your proposed security system

Qn.5d: Is the proposed security technology actually necessary?

The focus of this section is on whether a security technology is *necessary*, not whether that technology is *desirable* or *useful*.

Just because you possess the technical expertise to produce a security technology, that fact alone does not automatically mean society *needs* that technology. And if/when citizens and governments realise an unnecessary security technology is redundant then it is highly unlikely they are going to want to pay for it. However, technologies which are considered *necessary* by both governments and citizens to promote security and fight crime are those most likely to be implemented first, to be more likely to enjoy public support, and be the most resilient to budget cuts and similar constraints. Thus being able to identify the *necessary* from the merely *useful* or *desirable* should be at the forefront of any decision by designers when considering new projects and products. This is not an easy process as security agencies and governments may not be willing to divulge information on known security weaknesses. Also certain crimes/threats may be occurring which have not been identified yet, thus representing a problem we do not even know exists.

What makes this process more challenging is that security levels and specific crime rates have a temporal component; i.e., they fluctuate over time with the changing circumstances. Thus what is necessary today may not be necessary tomorrow.

Challenge: To identify security technologies which are *necessary* and/or those holes within our current security processes and crime detection/response procedures which *necessarily* require a technical solution.

Challenge: To produce a security technology that remains *necessary* when the security and crime threat levels fluctuate regardless of direction; i.e., both when they increase but also when they decrease.

Step 1. Begin by describing both your proposed security technology and the crime or security problem(s) it seeks to address.

Step 2. Using the information from Step 1 can you come up with convincing arguments as to; (a) why your proposed security technology is *necessary* to fix these identified problems?, and (b) why these identified problems *necessitate* a technological solution? For each 'yes' answer to (a) and (b) set out the reasons here. If either

is a 'no' explain here why this project should go ahead given the technology cannot be justified as necessary.

Step 3. Consider how your proposed security technology addresses the security/crime problem(s). Are there other technologies available which can also address these problems? If 'no' move to Step 4. If 'yes' compare your solution to the existing ones. Do the others achieve the same results in potentially a more acceptable and/or less intrusive/controversial manner than yours? If 'yes' then why should your proposed technology go ahead?

Note on Step 3: When comparing your proposed security technology to existing alternatives consider both how successful they are and how intrusive they are (i.e., how much information they collect, their effects on rights and liberties, whether they require physical contact, etc.). If they are sufficiently successful and less intrusive in their approach to your technology then you may have problems arguing yours is *necessary*.

Step 4. How can you design your technology such that it can respond to changing security/crime threat levels so as to remain *necessary* regardless of the situation?

Qn.5e: Will minority groups be bearing the security burden?

Certain security technologies (such as profiling technologies) are based entirely on identifying those individuals whose characteristic fit a certain predetermined image (or profile); often so as to facilitate secondary screening processes. Profiles are by necessity minority groups; we use profiling to whittle down all the subjects so as to hopefully identify only those of interest and thereby maximise the application of our resources – if the total number of subjects who matched the profile are too large (which depending on the population size can be much less than 50% of all those screened) then resources will not be available to conduct secondary screening and our profiles will have failed. The use of profiling, while potentially valuable from a security perspective, has evoked considerable negative social reactions and we need to understand the underlying issues.

Delve down into the characteristics of individuals and you soon realise that every single person in a society is a member of any number of minority groups; all it takes is the application of enough defining factors when grouping them to create a minority group, and you rarely need more than one or two factors to achieve this. It could be the realisation of this (that we are all minorities), or the fact that we do not always get to choose whether we belong to a group (consider medical conditions or ancestry), that has led society to shy away from, be wary of, and place restriction upon, security measures which deliberately target identifiable minorities.

When employing security technologies that can target minorities or place a greater burden on them, a number of negative resulting effects can arise:

- By deliberately choosing to place a greater burden on the rights and liberties of this group so that those in the majority may enjoy security benefits without paying a similar price we are discriminating against the minority group.
- The overwhelming number of those within the minority group will be false positives. These people will often face repeated secondary screening, especially when systems cannot learn from their mistakes.
- We may be placing individuals at greater risk of harm, abuse, and discrimination by including them in these minority groups.
- By designing security technologies which are not accessible to those within certain minority groups we can easily discriminate against them.

Our goal should be to design security systems which can effectively allocate resources and maximise security for all without increasing the burden for those in minority groups.

Challenge: To design security technologies which can accurately identify actual threats and criminals (a minority group), minimising the number of false positives and the negative effects of their mistaken identification, and provide equal security benefits for all while justly spreading any concomitant burdens to other rights and liberties.

Step 1. Explain how the proposed security technology differentiates between the

target group and the rest of the population?_____

_____

Step 2. Apply the explanation from Step 1 to the following groups. Could the proposed technology be configured such that members in minority groups based on the following characteristics would have a greater probability of being targeted as suspects?_____

□ Ethnicity □ Skin colour □ Religious affiliation
□ Political affiliation □ Age □ Sex
□ Physical/mental conditions

Step 3. What safeguards can you build into the technology to prevent the systematic discrimination of such groups whereby they bare the greater security burden?_____

_____

Note on Step 3: The key words here are *systematic discrimination*. A safeguard could be a system whereby individuals from these groups are only targeted if credible information of a specific threat exists and only then for the duration of that specific threat.

Step 4. To reduce the repetitive re-screening of false positives, will your technology be able to learn from past mistakes and not target these individuals in the future?_____ If 'no' go to step 5. If 'yes' how do you plan to achieve this?_____

_____ (go to Step 6)

Step 5. What is your plan then for dealing with false positives?_____

_____

Step 6. How can the system be adjusted to modify the false-positive/false-negative rates in response to specific information and threats?_____

_____

Step 7. Will a minority of the population be at greater risk of physical harm from the security technology?_____ If 'yes' how can the technology be designed to respond to this problem?_____

_____

Note on Step 7: Options might include: the security technology identifying these people and modifying its response accordingly; protections whereby the end-user is forced to check if the individual falls within these groups before using use; issuing warnings; etc.

Step 8. Will there be a minority of the population who will not be able to use the proposed security technology?_____ If 'yes' how can any discrimination, disruption and potential embarrassment to this group be mitigated?_____

_____

Qn.5f: Is the proposed security technology a disproportional response?

This section focusses solely on whether the proposed security technology represents a proportionate or disproportionate response to the security/crime problem it is built to address. It is not concerned with the necessity of that technology (see Qn.5d), its usefulness, or desirability.

While we all accept the need to respond to security and crime threats, if society considers your response disproportionate we risk undermining public trust and causing more damage than we have solved, and the negative effects of such disproportionate actions can remain in peoples' memories for years. For example, the use of batons, tear gas, and water cannons against peaceful civil rights demonstrators in the United States during the 1960's shaped a generation of Americans and remains a sensitive issue today. It damaged both race relations and the relationship between citizens and the state.

With today's security technologies and the advent of the digital world, disproportionality goes beyond just end-user responses and the effects on society. It includes issues such as excessive functionality, excessive usage, and the inability of technologies to distinguish between legitimate targets and innocent citizens. The potential of these pit-falls arising may be exacerbated by the constant pressure from governments and end-users on developers to produce technologies with ever greater functions, capabilities, and capacities, regardless of whether they intend to use these functions upon deployment.

There are also issues of whether the proposed security technology represents a disproportionate response to the problem it seeks to address, or whether use of the technology is perceived as excessive by the public.

Challenge: To design a security technology:

- That represents a proportional response to the security/crime problem it seeks to address;
- That only possesses the features/functions it needs to achieve this goal, while at the same time can be upgraded and downgraded so as to remain proportional as crime and security levels fluctuate;
- That does not undermine other rights and liberties of citizens.

Step 1. On the left side of a suitably large piece of paper, using short bullet-point descriptions list off all the specific security and crime problems that your proposed security technology is going to be designed to address (be brief for each and try to avoid grouping a range of offences/security-issues together with a single heading – instead list off all the individual crimes/security-issues as individual points).

Step 2. Now on the right hand side of the same piece of paper, using short descriptive bullet-points list off all the features of your proposed security technology (i.e., list off all the things that your proposed security technology can do). Again break them down into specific individual components rather than using larger generalised headings.

Step 3. Now draw lines linking the left-hand side *crime problem points* to the right-hand side *functionality points* where specific functionality points effectively address specific crime points (each point can link to as many other points as is relevant).

Step 4. Examine your page of now-linked problems and functionalities.

- Are there any (left-hand side) security/crime problems which are not linked to functionalities?_____ If 'yes' then your proposed design will not be able to address those security/crime problems.
- Are there any (right-hand side) functionalities which are not linked to specific crime/security problems?_____ If 'yes' then your proposed design has *redundant features*; are there any justifications for not removing these functionalities, and if so what are they?_____
- Ignoring any unlinked left- and right-hand side points, examine how the right-hand side functionality points are linked to the left-hand side security/crime points. Are there any right-hand side functionality points which could be erased such that all of the left-hand side security/crime points remain linked to at least one right-hand side functionality point?_____ If 'yes' these erasable right-hand side points represent *excessive functionalities*. Are there any justifications for not removing these functionalities, and if so what are they?_____

Step 5. Can your proposed security technology be designed to distinguish between offenders and innocent civilians?_____ If 'yes' are you building such features into the design of your technology?_____ If 'no', explain why not._____

Step 6. Reflect upon your proposed design and the security benefits you are trying to achieve. Can you honestly say this is the least intrusive method/approach/design for dealing achieving these benefits?_____ If 'no' how can you justify the design of this security technology, and are there any possible ways you can modify your design to be less intrusive while still achieving the same security goals?_____

Qn.5g: Will public support come with conditions attached?

Public support for all security technologies comes with conditions attached. At the most basic level all security technologies come with the condition that they are not systematically misused or used illegally: we support guns for police on the condition they do not start randomly shooting people on patrols, we support CCTV cameras on the condition they are not deliberately set up to look inside our bedroom windows; we support Tasers on the condition they are not used to torture immobilised cuffed prisoners; etc.

Some technologies have multiple bespoke conditions attached. For example airport whole body scanners come with conditions regarding who views the images, how they are viewed, that the images are not recorded, that the low radiation dose from the backscatter scanners is not a health risk, that selection processes are not based on racial or religious discrimination, and that those being scanned will be treated with respect.

Social acceptance of security technologies always depends on conditions being attached to them. By examining past security technologies and conditions for use, a reasonable conclusion is that the more controversial the security technology the more likely additional conditions will be attached to it. The challenge for those designing security technologies is that these conditions for support require the public knowing about the proposed security technology, its capabilities, and how/where it will be deployed and operated. As the public are often not included in the design process of a new security technology, the developers may find themselves rapidly trying to modify their products after release in the face of an angry public who are demanding both design and operation conditions be attached to the technology if it is to remain in use.

Challenge: To identify those aspects of the proposed security technology's design which are likely to result in conditions for support being attached to this technology. Then, having identified these design aspects, adding measures to allow these conditions to be met *before* the technology is released.

Step 1. Consider other security technologies or security procedures which already exist and perform either the same or similar operations as your proposed security technology. Identify and write down here all of the conditions that have been attached to these technologies governing how they are to be used.

Step 2. Using the information from Step 1, will the design of your technology be such that either these conditions will not apply to it or that your technology will be able to meet these conditions immediately upon release?_____ If 'no' justify here why you believe your technology will be considered acceptable without being able to meet the conditions of use that have been attached to other similar technologies/procedures?

Step 3. Are there any previous security technologies which did what your proposed security technology will do but failed to achieve public acceptance and were subsequently pulled?_____ If 'yes' what conditions were attached to these technologies and why did they fail regardless?

How can you design your technology to succeed where others failed?

Step 4. Now consider the design of your proposed security technology without your developer's hat on but as a citizen who may be subjected to this technology. What would you specifically *not* want this technology to be able to do?

What conditions would need to be attached to the use of this technology before you would accept its introduction?

And how could you modify the design of your proposed security technology so that it could prevent these actions and meet these conditions?

Qn.5h: Will this security technology meet general standards of what is considered socially acceptable?

If security technologies are not socially acceptable they will not survive unscathed over the mid- to long-term. At best a developer will be forced to modify their designs in the face of negative public responses. They may also find restrictions place on the sale and use of this technology. In the worst case the technology will be banned from a society altogether. Given that such outcomes all affect the profits, credibility, and trust of the developer, there are real incentives in ensuring a security technology meets society's standards for acceptability from the moment it is released.

The problem is that a *society* is not a homogenous entity; different people and groups within society have different beliefs and views on what is and is not acceptable. That being said, by examining previous examples of security technologies which have elicited negative reactions you begin to see commonalities within the social responses. A number of commonly held values begin to emerge which you can use to guide your design choices and hopefully minimise the risk of social rejection.

Step 1. Consider your proposed security technology. In the past, has this type of technology proven unacceptable to society?_____ If 'yes' why do you believe your particular security technology will fare differently?_____

_____

_____

Step 2. Can you imagine scenarios whereby the proposed security technology would be considered unacceptable if; used in certain ways?, used against certain targets?, or used in certain situations?_____ If 'yes' how can you modify the design to prevent or minimise such usage?_____

_____

_____**(see also Qn.5g)**

Step 3. In the past has this type of security technology been perceived as fair?_____ If 'no' why is your proposed design going to fare differently?_____

_____

_____

Note on Step 3: While *fairness* seems very hard to judge at first, you can get an idea on what to include here by looking at previous examples of *unfair* technologies:

- The UK's failed National Identity Register with its cradle-to-grave collection of information on citizens.
- Mandatory biometric systems which people are forced to use to access services but which cannot cope with anomalies such as missing fingers, job-related wear, age, and a variety of medical conditions.
- Mosquitos which target young people regardless of their actions.

Step 4. Will your proposed security technology create new potential security risks?_____ If 'yes' does the public bear the burden of these risks?_____

Note on Step 4: For example the defunct National Identity Scheme proposed the *mandatory* collection of individual personal data to be held in central databases. This would become a prime target for attackers; hence the risks will be borne by the public.

Step 5. Could the proposed security technology be used to discriminate?_____ If 'yes' how can your design prevent this?_____

_____**(see also Qn.3f)**

Note on Step 5: Discrimination can be indirect; for example a profiling or data mining programme whose underlying data sets are themselves discriminatory will perpetuate this discrimination regardless of the best intentions of the end-user.

Step 6. Could this technology be used in such a way that it has a negative impact upon social cohesion?_____ If 'yes' how can your design protect against this?_____

_____

Note on Step 6: Such as comprehensive CCTV systems being deployed in residential areas selected because of many of the residents are Muslim.

Step 7. Could it be used to target vulnerable members of society, and/or treat the symptoms of social problems while ignoring the underlying causes?_____

Note on Step 7: Such as mosquitos marketed as a *solution* to people sleeping rough.

Step 8. Has this or similar security technologies been restricted or banned after social action in the past?_____ If 'yes' why did this happen and how will yours be different?_____**(see also Qn.3a)**

Step 9. Have there been incidents of misuse or abuse of similar technologies by end-users/officials?_____ If 'yes' how can your technology prevent such misuse?_____

_____**(see also Qn.7a)**

section 6:

issues of functionality

Qn.6a: Could the proposed functionality of the ST be open to criticism?

Functionality is defined here as *all the things your technology can do and how it does them*. It represents the culmination of all the individual *functions* (i.e. the individual components/abilities/modes-of-use/capabilities/etc.) you choose to build into your security technology.

The number of functions different security technologies possess varies hugely. At one end of the scale a technology may have only one intended function (e.g. a police baton – for hitting things), while at the other end a security technology may possess a vast range of intended functions (e.g. a national identity register – a multitude of data mining and matching capabilities, authentication capabilities, possible biometrics, linked to the access of state services, usable by different government departments for their individual bespoke needs, etc.).

You are probably not a charity, rather you wish to sell your proposed security technology once you have built it and make a profit when doing so. As such two temptations exist when it comes to your design. One is to *maximise its functionality* thereby increasing both (a) the range of potential purchasers and (b) its attractiveness to these potential purchasers. The other is to *exclude certain functions* from your design which may not be essential to its operation on the basis of reducing production costs, which in turn may also increase the range of potential purchasers, and its attractiveness to them.

Caution must be exercised at this point. Just because your security technology can be designed in a particular way does not necessarily mean you should design it that way, especially if you have not taken into account the risk of social resistance. Social resistance to security technologies in the past has arisen in response to functionality issues for three main reasons:

1. Unnecessary functions being included in the design of the security technology.
2. When necessary functions achieve their goals through methods which the public consider to be excessively heavy-handed or disproportional.
3. When the security technology lacks certain functions which the public consider essential.

Challenge: To achieve a design that will maximise the potential sales and profitability of your security technology. This goal requires that you minimise the risk your design will alienate the public and result in the sale of your technology being restricted or banned over the mid-term. This means you need to include the right mix and form of functions to avoid the three functionality issues listed above.

Step 1. Create a list of all the security/crime related outcomes you are specifically seeking to achieve by developing your proposed security technology.

Step 2. Now create a second list of all the individual functions your proposed technology would need to possess if it was to achieve the security/crime outcomes listed in Step 1 (only include those functions which are *essential* to achieving these outcomes).

Step 3. By applying your research and knowledge, are you aware of any existing security technologies similar to yours where the public demanded certain functions be incorporated into the design so as to avoid or respond to controversies related to its operation? If 'yes' list these functions off here.

Step 4. By combining the lists created in Steps 2 & 3 you have created a list of the probable essential functionalities for your proposed security technology. Create a list here of all the functions you intend to include in your design.

How does the combined list from Steps 2 & 3 compare to what you are intending to include in your design? (For those functions you had not included, how can you justify not including them now? For those functions you had included but were not on this combined list, how can you justify retaining them now?)

Step 5. Create a list of functions you intend to include in your design having now taken into account the output from Step 4. Could any of these functions possibly be considered heavy-handed or disproportional? If 'yes' can you modify these to achieve the same goals but in a more proportional manner?

Note on Step 5: To assist in answering the first part of Step 5 apply the following question: *Would I be happy for my loved-ones (i.e. wife/husband/parents/ grandparents/children/grand-children, etc.) to be subjected to this security technology in its current form?*

Qn.6b: What is the potential for function creep and dual-use?

Function creep is defined by Collins English Dictionary as "*the gradual widening of the use of a technology or system beyond the purpose for which it was originally intended*". Function creep also has the potential to feed into the concept of *dual-use*; the idea that materials, hardware, and knowledge with peaceful applications can be exploited for illicit ends. For our purposes the major difference between function creep and dual-use is that function creep needn't result in a security technology being used for illicit purposes (i.e. the new uses for this technology may be completely legitimate) whereas dual-use always results in the security technology being used for illicit purposes.

Dual-use, as defined above, is a relatively straightforward scenario for us in that no responsible developer (such as yourself) designs a security technology with the intention that it be used for illicit purposes, for they know such illicit use is always going to result in controversy and could severely damage their own reputation! Both the nature of this misuse and the number of times it occurs will influence the public's reaction to this security technology; amd we can expect this reaction to be quite sudden and intense, especially if the illicit use becomes the focus of national media reports, political debate, and judicial actions all in a short space of time. You should anticipate a robust official response through the introduction of sanctions, such as; disciplinary/criminal actions against end-users, tighter regulation and monitoring of future use, forced changes to design, limiting or even banning of sale/use, etc.

Function creep is less straightforward in that there can be many reasons why a security technology should be used for legitimate purposes beyond those originally intended. Also the public may call for and support this new extended functionality. However, the risk remains the public may not be willing to accept these new uses, especially if they only gave their acceptance to the technology being deployed in the first place on the understanding that it would not be used beyond its original purpose/functionality. There may also be pressure from the state/end-users to extend the functionality of a security technology to respond to perceived new threats and/or to improve the cost effectiveness of the device. If the public do not support this extended use there is a risk they may remove support from the technology entirely.

Challenges:

- To design a technology where the propensity for dual-use is minimised.
- To design a technology where function creep can be controlled so as to allow that which is socially acceptable but to restrict/prevent that which does not enjoy public support.

Step 1. Imagine that you are no longer a responsible developer of security technologies; rather you have *gone rogue* and are seeking to use your proposed security technology to commit criminal offences. List off the different ways you could use you security technology to conduct (or assist in the conducting of) different criminal acts.

Step 2. How can you modify the design of your proposed security technology to minimise or mitigate the use of your technology to perform the criminal acts you identified in Step 1?

Step 3. Create a list of all the individual functions and uses of your proposed technology.

Note on Step 3: As a reminder, *functions* are defined as all the individual components/abilities/modes-of-use/capabilities/etc. you choose to build into your security technology.

Step 4. Examine the list created in Step 3. From a social acceptability perspective are any of these functions/uses *mandatory* (in that public support is conditional upon the existence of this function/use)? If 'yes' then how will your security technology be designed such that these functions/uses cannot be altered or removed by the end-user upon deployment of the technology?

Step 5. Again examine the list created in Step 3, this time focussing on functions/uses that are *not included* in your list. From a social acceptability perspective are any of these omissions *mandatory* (such that public support is conditional upon the absence of this function/use)? If 'yes' how will your security technology be designed such that these functions/uses cannot be added by the end-user upon deployment of the technology?

Qn.6c: Does this security technology impose a burden on somebody?

If this Section applies you should also consider the applicability of Sections 2g & 4h.

While a security technology can be designed to benefit people (hopefully by providing them with security), its design also has the potential to impose a burden on somebody. Social resistance to a security technology can arise if this burden is not equally spread throughout all members of society. Leaving aside the legal issues surrounding possible discrimination, as well as the acceptability issue of what society thinks is an appropriate security response within a fair and democratic society, there are also the physical and temporal burdens imposed upon an individual by a security technology (let us call these *functionality burdens*). If these functionality burdens are not equally dispersed throughout all the subsets of society, a security technology runs the risk of social resistance forming, especially within those subsets disproportionately affected.

The functionality of your proposed security technology (i.e. all the things your technology can do and how it does them) not only determines how your technology will work, but also how it will physically interact with the public. This interaction component (i.e. the type of interface you design, what you expect people to possess or be able to do to use the technology, how long people are detained for processing by the technology etc.) has the corollary effect of determining, amongst other things:

- Who can and who cannot physically use the technology;
- How long it will subsequently take someone to undergo security screening;
- Whether or not an individual will be able to access services as a result of their ability/inability to use the security technology;
- Whether or not an individual will be materially disadvantaged by their ability/inability to use the security technology.

While all of these effects have the potential to create social resistance, developers should be particularly concerned when an individual is repeatedly subjected to such functionality burdens despite not having done anything wrong and through no fault of their own (i.e. the repeat false-positive). An example being a citizen with Parkinson's disease being unable to undertake an airport whole body scan because they cannot remain still, and as a result being subjected to more intrusive secondary screening. The failure by developers to at least attempt to mitigate the burdens imposed upon these individuals opens the door to social resistance led by the directly affected sub-groups.

Step 1. Create a list of all the individual things people need to be able to do and/or possess to use your proposed security technology or to be successfully *cleared* by your technology.

Note on Step 1: For example: stand completely still for 7-8 seconds; remember a password; have fingerprints; have all their fingers; not possess the combination of characteristics you identified as *possibly threatening* when you created your profile; not have shrapnel or metal pins within their body; etc.

Step 2. Now for each of the items on your list from Step 1 identify any groups within society who will not be able to meet these requirements.

Note on Step 2: I have used 'groups' and not 'individuals' within the wording of Step 2 as there will almost always be more than one person who will not be able to meet each identified item. Examples here may include: amputees; workers within certain industries; Parkinson's disease sufferers; Alzheimer's sufferers; the elderly; religious affiliations; etc.

Step 3. Now list off all the consequences for the groups identified in Step 2 for being unable to meet the requirements identified in Step 1.

Note on Step 3: For example: not being able to fly; more intrusive secondary screening; not being employable in certain role; time delays; humiliation; more intrusive questioning and subsequent lack of privacy; resentment; etc.

Step 4. Looking at your list from Step 2 and the consequences identified in Step 3 ask yourself the question; *Do these people deserve to be treated as potential security threats suffering the identified consequences because of their membership of their respective groups?* If 'no' then you will probably need to either modify the design of your technology, mitigate the effects on these groups, and/or provide alternatives for them if you wish to avoid controversy/resistance, for you can be sure the members of these groups (and the wider society) probably feel the same as you do when answering this question.

Step 5. How can you minimise the effects of your proposed security technology to those who cannot use it, as well as those false-positives who will be repeatedly targeted?

Note on Step 5: For example: alternative screening processes, pre-screening registration; automatically updating databases, identification systems for certain groups, end-user training; enabling end-users to override automated systems, etc.

Qn.6d: Is your security technology susceptible to excessive and/or repeating errors?

There is no such thing as 100% security. The perfect, infallible security system simply does not exist. As such errors will always result from the use of any security technology in the form of *false-positives* (innocent individuals wrongly identified as a threat) and *false-negatives* (attackers wrongly identified as non-threatening) or of people simply being missed altogether. This is why security experts recommend the adoption of *layered-security*; a range of different overlapping security measures, so as to minimise the number of those who slip through the security net, while hopefully not repeatedly targeting the false-positives.

However, just because there will always be false-positives/negatives this does not mean you 'the security technology developer' can afford to ignore these problems when designing security technologies. Security technologies which repeatedly produce unacceptably high levels of false-positives/negatives run the risk of alienating both end-users (who will come to view the technology at best as unreliable, or at worst as a broken waste of time that should be ignored) and the public (who may come to resent being subjected to the technology, will question its cost and deployment, and may result in a loss of trust in the end-users, officials, politicians, etc.).

Challenge: The challenge for developers is to design a security technology that can both control the rate at which errors will occur, as well as efficiently address the errors and their effects at the point when they do occur. This requires the design of a technology which:

- Controls the rates of false-positives/negatives;
- Recognises where errors occur within the technology;
- Validates the accuracy of data that is inputted;
- Allows the data which produced any decisions made by the security technology to be questioned by end-users;
- Allows the results of the technology to be questioned by end-users.

Step 1. (If it is possible) are you including within your design the ability to adjust the rates of false-positives and false-negatives?_____ If 'no' (and it is possible) how can you justify this choice?_____

_____

_____

_____

Note on Step 1: False-positive and false-negative rates are linked, in that increase the rate of one will likely decrease the rate of the other, and *vice versa*.

Step 2. Does your proposed security technology at any point rely upon data inputted by humans (or machine inputted data where that data was originally collected and inputted by humans)?_____ If 'yes' what system do you have in place to double check the data so as to remove/reduce human input error?_____

Step 3. Will the reliability of your proposed security technology fall below acceptable levels as the amount of data inputted increases?_____ If 'yes' how are you planning on dealing with this problem?_____

Note on Step 3: For example; within a one-to-many-matching identification system with a 99.99% accuracy rate and a database of 1,000,000 people, every individual would be falsely identified as 100 other people every single time they are scanned. This system would be useless and quickly ignored by end-users due to its unreliability.

Step 4. Will the end-user be able to see and review the data used to make an automated decision so as to question its accuracy?_____ If 'no' why not? If 'yes' how will you ensure the data is prevented in a format allowing for a meaningful review?_____

Step 5. Will the output (including any decisions made) of your proposed security technology be open to questioning by the end-user, and open to being overridden by the end-user?_____ If 'no' why not?_____

Will the end-user be able to modify the data held by the security technology in response to a false-positive so as to prevent and individual repeatedly being mistakenly identified as a security threat?_____ If 'no' why not? If 'yes' how will this process occur?_____

Qn.6e: Will the proposed security technology be capable of, and effective at, addressing the identified security problem(s)?

It is all well and good to build a security technology with multiple functions, which is cost effective, does not pose a health risk, meets all legal and human rights criteria, is not open to end-user abuse, and does not create additional safety and security issues. However, such a design will all be for nothing if that security technology is either incapable of, or ineffective at, achieving its ultimate goal; i.e., tackling the security problem(s) it was created to address. Failure to achieve this end will result in social resistance as the public become aware of this fundamental shortcoming, for the public resent paying for something which doesn't work and being subjected to a security measure which does not provide security. This failure affords ammunition to both the media and the political opponents of those who authorised the purchase and introduction of such a technology, opening the door to a potential loss of public trust in, and respect for, those government agencies and the end-users.

Challenge: When seeking to demonstrate that your proposed security technology is both capable of, and effective at, addressing the identified security problems, the initial question you will probably ask is '*will your proposed security technology actually work?*'. While this is a good starting position it is a question where a 'yes' answer can also paper over any number of fundamental cracks. You need to go deeper in you examination by asking (and answering) the tough underlying questions if you wish to demonstrate the security-providing pedigree of your proposed technology. Other questions to be answered include the following:

- When you say it 'works' are you actually saying it 'meets minimum standards?'
- While it may work, will it be effective?
- And if it is effective, will it be consistently so?
- And assuming your technology is consistently effective, will it only target people engaged in criminal activities or will everybody be effected?
- Finally, will your technology address the root (or cause) of the problem, or is it designed only to address the symptoms?

Step 1. *Will your proposed security technology work?* To answer this question first define what *working* is, framing your answer as a list of one or more measurable values (such as; an *n*% reduction in a specific crime; an *n*% increase in arrests or seized goods, etc.) where current values exist so as to permit a comparison.

What experiments/pilot-studies/etc., have you carried out or are planning on carrying out so as to measure whether or not your proposed security technology achieves the criteria/criterion you designated as defining *working*?

At the point in the design/validation process where you can compare the values from your experiments/pilot-studies/etc. against your measures for success, does your security technology qualify as capable of addressing its targeted security problem(s)?_____ If 'no' how are you going to respond to this situation?_____

Step 2. How do you plan to demonstrate that your proposed security technology works consistently, regardless of:

a) Time in operation:_____

b) The different circumstances under which it is deployed:_____

Note on Step 2: To prove the *consistent* operation of your proposed security technology you are going to have to demonstrate that it remains effective for a sufficient time period beyond the moment of deployment. Also you will need to determine those situations/environments it operates effectively in and those it does not (if any).

Step 3. Does your proposed security technology effectively differentiate between those citizens acting illegally and those not, such that it only affects the former which ignoring the latter?_____ If 'no' justify why your proposed security technology should be deployed given its inability to discriminate between offenders and non-offenders?_____

Step 4. Will your proposed security technology be able to effectively reduce the incidents of specific crimes by preventing/deterring people from committing these crimes, or does it just focus on catching offenders after they have committed these crimes?_____ If it is the latter how can you justify the effectiveness of your security technology?_____

Qn.6f: Is the technology underpinned by scientific and/or operational uncertainty?

To produce a completely new security technology you may not need to push the boundaries of science or engineering. If the technologies you are relying upon are well developed and understood then any uncertainty as to the operation, safety, and effectiveness of your new security technology will be minimised. Equally however, producing a new technology may require you to either utilise existing technologies in a novel manner and/or to push the boundaries of current science by developing something completely new. In these circumstances any uncertainty inherent to your final product may have a role to play when it comes to the social acceptability of your technology.

If you are pushing the boundaries of what has been achieved before, questions will always arise as to the effectiveness, safety, and acceptability of your product. In the past, security technologies have suffered as a result of uncertainty in different forms:

- Backscatter whole body scanners and various less lethal weapons faced questions over their safety. For whole body scanners this has led to lengthy delays on their widespread introduction as the health issues are debated.
- The UK's national identity scheme, mass biometric systems, and various data mining systems have all faced resistance as the result of uncertainty over their potential effectiveness.

The reason why *uncertainty* should not be ignored is that it is the perfect tool for assisting the resistance of a technology. It is very hard to argue that a new technology is perfectly safe, reliable, and/or effective, if the evidence from previous use is not there to back these arguments up. In these circumstances those against the introduction of a particular security technology do not need to present any evidence themselves justifying their position. They can merely rely upon the absence of certainty to justify and validate their position.

Challenge: To have accumulated enough evidence through testing to reduce any uncertainty surrounding your proposed security technology *before* it is released for sale and/or implementation. Also to be in a position to present this evidence in a form which permits independent assessment at the moment the technology is released.

Step 1. Are you planning on building something where some or all of the components are sufficiently innovative to be considered *new*, or are you only utilising existing technologies which are well understood in your final product design?_____

_____

_____

Note on Step 1: By *new* I mean either something which has never been developed before, hence you are uncertain as to its final properties/how it works/etc., *or* the novel application of an existing technology, again where you are uncertain as to its final properties/how it works/etc. You will need to apply your own judgement here.

If you are using something new (as determined in Step 1) then continue with the Steps below.

Step 2. Does the technology pose a potential risk to the health of the target, no matter how minute?_____ If 'yes' what steps have you taken (or are going to take) to prove the safety of your technology, including any independent verification of this?_____

Step 3. Given your technology is *new* how are you going to prove the effectiveness of your proposed security technology?_____

Note on Step 3: For example: trials in different operational situations and over extended time periods; independent verification of results; engagement with critics and/or concerned social groups; etcetera.

Step 4. How are you planning on engaging with the public to present the evidence you have collected on your proposed security technology so as to reduce the effects of any inherent uncertainty?_____

section 7:

safety, security, misuse, & abuse

Qn.7a: Is there a history of abuse or misuse of similar security technologies?

A large number of security technologies have, in the past, been abused or misused. The scope of this abuse ranges from the mischievous (e.g. tax officials illegally accessing databases to look up the tax details of their neighbours or celebrities out of curiosity) to the catastrophic (e.g. groups, governments, and/or foreign invaders using identity-card information to perpetrate genocide against religious and/or ethnic groups). Those carrying out this abuse/misuse ranges from governments to state agencies, individual end-users, insiders, commercial organisations, citizens, and external attackers.

As the developer of security technologies you can either choose to ignore any historical abuse of security technologies similar to your proposed design, or you can take the time to examine when and how these technologies were abused so as to inform your own design and hopefully produce something which is less susceptible to such abuse. There are benefits to adopting the latter approach.

Social resistance to security technologies in response to past abuse can be widespread and long-lasting, resulting in a restricted market for your final products. Various countries have developed strict laws regarding the use of security technologies (especially in relation to privacy) in response to the past abuses by oppressive regimes. Citizens in some of these countries are also less accepting of new security technologies. In Germany, Spain, and various South American countries (all of which have suffered under various regimes) there is less acceptance of new security technologies and greater restrictions placed on technologies which impact the privacy of the citizen (such as CCTV, airport whole body scanners, ID cards, etc.).

On the other hand by developing your proposed security technology so as to avoid the abuse/misuse pitfalls of competitors you have the opportunity to benefit from any social resistance to the alternative designs. You can rightly market your design as the safer, socially acceptable option, while highlighting to politicians and state agencies the potential reputational benefits they could reap by introducing your technology.

Step 1. Start by identifying and listing below instances of where security technologies similar to your proposed technology have resulted in controversy resulting from their abuse or misuse.

Using these identified instances, I want you to distil them down so as to create a list of design specifications which capture all of the fundamental design characteristic which

either created the controversies, underpinned the controversies, or allowed them to occur.

Note on Step 1: For example, for whole body scanners this list may include:

- Created images which display anatomical details including genitalia;
- Incorporating the ability to store images;
- Having the scanned images visible to the scanner operator; etcetera.

Step 2. Using the list of design characteristics from Step 1, create a list of alternative characteristics which address these issues that a future security technology could incorporate so as to avoid controversy.

Are you going to incorporate these alternative characteristics into your own design?_____ If 'no' justify why you are taking this decision.

Step 3. What are your plans for setting yourself apart from these previous controversial technologies so that you are not tainted by association?

Step 4. How are you going to convince the public that your technology will not be abused or misused given that similar technologies have been in the past?

Qn.7b: Could the *operation* of the security technology potentially jeopardise the safety and security of citizens?

It goes without saying that security technologies are designed to provide citizens with some additional measure of security. However it is sometimes the case that the way in which certain security technologies operate (i.e., how they work to achieve their security goals) can have the opposite effect by jeopardising the safety and security of citizens. Obviously this is not the developer's intention. When the operation of a security technology can be shown to jeopardise the safety and security of citizens, not only will this lead to resistance from those directly put at risk, but it opens the door to wider social resistance should those not directly affected take up the cause of those who are and remove their support for the technology as a result.

Usually those who are jeopardised by the operation of a security technology constitute a small minority of the population (if it was the majority then one hopes the developers would realise this and be smart enough not to try and release that technology in its current form). The fact these individuals are placed at risk at all usually comes about for any of the following reasons:

- Those adversely affected form a very small minority group(s) of individuals sharing one or more particular traits which make them particularly susceptible. Unless the designers were specifically aware of this group's existence and the threats they face then the fact their concerns were not contemplated is understandable.
- The safety/security threats to individuals were unrealised, unforeseen (potentially unforeseeable) consequences of the security technology's operation.
- The designers were focussing on the overall security benefits and felt that any concomitant risks were justifiable.
- The designers assessed the safety of their technology by assuming it is being operated in the manner they intended; i.e., they are assuming the technology will not be abuse, misused, or used for non-intended purposes.

Challenge: To force yourself to look beyond both the *obvious* and the *intended* when you are designing your security technology and visualising how it will work. Hopefully this will assist you in reducing the propensity for your proposed technology to place at risk the safety and security of individuals through its operation.

Step 1. Imagine you have completed your proposed security technology. It has been deployed in the field and operates in the manner you think it should. Can you think of any individuals/groups who despite not engaging in any illegal activities would not want to be subjected to your technology out of fear that by doing so their safety/security is placed at risk?_____ If 'yes' list these individuals/groups here along with the reasons behind their fear._____

Note on Step 1: For example, the proposed UK's National Identity Register (now destroyed) would have recorded information on every UK citizen including their current address, with a legal requirement placed on the individual to ensure this information was kept up-to-date. However, there are many legitimate reasons for not disclosing your location; women fleeing domestic violence or forced marriages, criminal informants, witnesses in criminal cases, etcetera. These groups would have been placed at risk by the operation of this security technology, and yet were not obvious categories of concern.

Step 2. Can you think of any circumstances where the successful operation of the security technology forces individuals or groups to modify their behaviour such that they must take on a greater level of personal risk to their safety/security? If 'yes' list these individuals/groups off here and describe the circumstances.

Note on Step 2: For example, the use of Mosquitos in public places forces young people to disperse, often to areas with less pedestrian traffic. Yet young people themselves report that one of the main reasons they congregate in public spaces to begin with is that they feel safer there because there are more people about. This technology therefore forces them to leave public areas and move to places they feel less safe in.

Step 3. Going back to the approach used in Step 1, again imagine you have completed your proposed security technology and it has been deployed in the field, only this time it is being used for purposes other than you originally intended. Can you think of any lawful ways your proposed technology could be used such that the safety or security of individuals/groups (who are not doing anything illegal) could be placed at risk?
If 'yes' describe these unintended methods of use and how they place the individuals/groups at risk.

Use the answers from these three Steps to inform the design of your proposed security technology by converting any affirmative answers into design specifications.

Qn.7c: Could the *presence* of the security technology potentially jeopardise the safety and security of citizens?

In a similar vein to Qn.7b, while security technologies are designed to provide citizens with security, sometimes the *presence* of a particular security technology can jeopardise the safety and security of citizens. Obviously this is never the developer's intention.

All security technologies have *ripple effects*. Initially they may reduce a particular crime and/or increase prosecutions. They may also alter the offending patterns of criminals; i.e., criminals may change where they commit offences as well as the types of offences they commit. Additionally technologies may impose costs on citizens through higher fees and taxes or time-costs. These (and more) are all unintended consequences that ripple out from the introduction of a security technology. To have a chance at success, the benefits brought by a security technology need to outweigh the negative effects.

One of the unintended ripples of any security technology is that its presence has the potential to jeopardise the safety and security of citizens. This not uncommon as security experts agree that most security technologies create new security problems. As such you should not reject a proposed security technology just because you can identify situations where its presence could produce harms. However at the same time this does not mean that you can ignore such effects.

You can quickly lose public trust and support when an unacceptable number of citizens are harmed by the presence of a security technology, or where both the risk of harm and the potential consequences of that harm occurring are considered too high. By identifying the potential safety and security risks brought about by your proposed technology you will afford yourself the opportunity to address these issues during the design process. Hopefully this will minimise both the size of any risk caused and the number of people who will be forced to bear that risk.

Step 1. In this Step you get to play the role of an attacker. Assume your proposed security technology has been built and deployed (congratulations!). As the attacker, is there value to be had from attacking the security technology? In other words, is the security technology a target worth attacking such that the potential gains from doing so outweigh the risks?_____ If 'yes' how would you carry out attacks on your own proposed security technology?; why these attacks?; and what do you predict the likelihood of success/failure/getting-caught to be for each of these attacks?_____

Step 2. If an attacker successfully attacked your proposed security technology, what other (new) crimes would they now be able to carry out as a result of this successful attack?

Note on Steps 1 & 2: Could your security technology become a potential *honey-pot* for attackers? For example, if your technology stored large amounts of personal information on all citizens in a country in one place then it would become a prime target for criminals and state-sponsored hackers/attackers.

Step 3. Using the attack methods you identified in step 1, can you modify your technology so as to reduce the possibility and consequences of a successful attack? If 'yes' how would you achieve this?

Step 4. How likely is it that your security technology could be under attack (or was attacked in the past) and that you would not be aware of this?

How could you modify the design of your proposed technology so as to make attacks (both past and present) more recognisable?

Step 5. Will your proposed security technology be the most *resource efficient* solution available to tackling its intended security problem? If 'no' justify why it should still be introduced.

Note on Step 5: If your proposed security technology diverts or ties-up security resources (money, people, time, etc.) away from more efficient/effective technologies and/or more important/larger threats, then the presence of your technology could be jeopardising the safety and security of citizens.

Qn.7d: Could *design shortcomings* of the security technology or the *lack of built-in protections* potentially jeopardise the safety and security of citizens?

Not every security technology which is designed, developed, and deployed works perfectly first time, every time. Given that it is impossible to test a security technology against every possible eventuality that could ever be conceived it is highly likely that some safety/security problems with a technology will only become evident (or even discoverable) after the technology has been deployed in an operational setting. Furthermore some safety/security issues may have been considered so unlikely to arise or be exploited that the developers did not consider these necessary to address.

However, given the number of security technologies which have been misused, abused, avoided, fooled, and/or circumvented in the past, all leading to the safety and security of citizens being placed at risk, there is no excuse for the designers and developers of security technologies not to assume their technology will be attacked, misused, and/or abused once deployed. To assume otherwise is to ignore reality and in turn will open the door to criticism and ridicule by those citizens adversely affected by the failing of a security technology.

You should never assume your proposed security technology: will operate perfectly; in all situations; exactly as intended; and will provide perfect security. By accepting the inevitable failings of your future technology you will be in a better position to critically and honestly assess your design and respond to the flaws you see there.

You must also assume that a minority of end-users will abuse your security technology. Given the negative social response to such abuse, and the probability of damaging media reports arising from such abuse, it is highly recommended you assume such abuse such that you can design against it.

Challenges:

- To recognise the value of your proposed security technology while at the same time being willing to identify those shortcomings within your proposed design that will have the potential to jeopardise the safety or security of citizens.
- To recognise the value of your proposed security technology while at the same time recognising that some end-users of your technology will abuse it in the future.

By acknowledging the truth of these two statements you will hopefully be better placed to actively address them as far as is possible through the design choices you take.

Step 1. Your proposed security technology has been deployed. If you were an attacker, how easy would it be for you to avoid the security technology? (Describe the different ways you could continue to carry out those illegal acts that your proposed security technology is attempting to address *without ever coming into contact with the*

technology)_____

_____

_____

_____

Step 2. Again acting as the attacker; if you did have to come into contact with your proposed security technology, describe the different ways you could fool the security technology into treating you as a false-negative, thereby allowing you to continue carrying out the illegal acts your proposed security technology is attempting to address?_____

_____

_____

_____

Step 3. Does your proposed security technology include a built-in oversight capacity or function so that the actions of end-users can be monitored in real-time so as to prevent or deter any illegal use of the security technology?_____ If 'no' explain why not. If 'yes' describe how this works._____

_____

_____

_____

_____

_____

Step 4. Does your proposed security technology record information on end-user use so as to deter/prevent/prosecute any illegal end-user use?_____ If 'no' explain why not. If 'yes' describe how this works._____

_____

_____

_____

_____

_____

Step 5. Does your proposed security technology provide the client with sufficient information/feedback so that they can determine whether somebody is misusing or abusing the technology, either in *real*-time or as soon as possible after the abuse has occurred?____ If 'yes' explain both how it achieves this and what information it provides the client. If 'no' explain why not._____

_____

_____

_____

_____

_____

Qn.7e: Is there a propensity for the security technology to be abused or misused thereby jeopardising the safety and security of citizens?

Security technologies are designed to enhance the safety or security of citizens on the assumption they are used in the manner the designers anticipate, and that they are not abused. However both those who commission security technologies and those entrusted with their use (i.e., governments, government agencies, commercial businesses, public and private employees, police and other state officials, etc.) have a history of misusing and abusing these technologies. Sometimes this is at a state level while at other times this represents the actions of rogue individuals. When a security technology is abused, the public can lose trust in both the end-users and the technology itself.

Those charged with using a security technology are uniquely placed to abuse that technology. Often the use of security technologies (or the use of security technologies in specific ways) is restricted to selected groups of people; these groups are trusted and considered reliable guardians of the technology because of their position and training. Unfortunately because these groups are trusted and/or because of a lack of resources, any oversight functions (whereby the watchers are watched) are not always robust. The result being the door is often open for end-users to abuse their security technologies. This abuse can arise from a variety of situations, including:

- People can give in to temptation to satisfy their curiosity, out of impulse, or if they see an opportunity.
- The end-user may have been corrupt all along.
- The end-user may be stressed, emotional, angry, scared, etc., when using the security technology and thus do not act as they would if they were calm/rational.
- The end-user may be being coerced or bribed into acting illegally.

If your proposed security technology has the potential to jeopardise the safety and security of citizens if abused or misused (and given the fact that many existing security technologies have been abused or misused by end-users in the past) then you must work on the assumption that end-users of your proposed security technology will attempt to abuse your technology. As you are obviously a conscientious designer/developer who does not want to see people harmed by the misuse or abuse of your technology then the onus falls upon you to identify how your technology could be abused, and by whom, so that measures can be taken to prevent (or at least minimise) this misuse/abuse.

Step 1. Identify here as many of the different ways your proposed security technology could be abused or misused as you can.

Step 2. Now identify which of the following groups could abuse your proposed security technology in the ways you identified in Step 1. For each group you identify write down how you believe this groups could be prevented from carrying out this abuse.

- Governments: ;
- Police: ;
- Other law enforcement officials (identify which): ;
- Other public servants (identify which): ;
- Private contractors: ;
- Businesses: ;
- Citizens: ;
- Foreign powers: ;
- Others (identify which): ;

Note on Step 2: Where possible try to identify design opportunities for your proposed security technology when identifying methods for preventing/minimising abuse and misuse by the different actors above.

identified design opportunities

Design requirement:_____

Section:___ Qn:___ Step:___

Details:_____

Design requirement:_____

Section:___ Qn:___ Step:___

Details:_____

Design requirement:_____

Section:___ Qn:___ Step:___

Details:_____

Design requirement:_____

Section:___ Qn:___ Step:___

Details:_____

Design requirement:_____

Section:___ Qn:___ Step:___

Details:_____

Design requirement:_____

Section:___ Qn:___ Step:___

Details:_____

Design requirement:_____

Section:___ Qn:___ Step:___

Details:_____

Design requirement:_____

Section:___ Qn:___ Step:___

Details:_____

Section Individual Question	Design Opportunity	Policy or public-engagement opportunity	Future work required
1. PHYSICAL OR MENTAL HARM			
1a. Health risks outweigh the benefits	☐	☐	☐
1b. Health risks unacceptable to the public	☐	☐	☐
1c. Health risk is greater for 'at-risk' groups	☐	☐	☐
1d. The ST causes pain or discomfort	☐	☐	☐
1e. The ST causes physical injuries	☐	☐	☐
1f. The ST causes a fatality	☐	☐	☐
1g. Failure to acknowledge the health risks	☐	☐	☐
2. LIBERTIES & HUMAN RIGHTS			
2a. The ST impacts ones right to privacy	☐	☐	☐
2b. No respect for informational self-determination	☐	☐	☐
2c. Right to assembly and association impacted	☐	☐	☐
2d. Freedom of expression impacted	☐	☐	☐
2e. Right to a fair trial impacted	☐	☐	☐
2f. Losses to rights outweigh security benefits	☐	☐	☐
2g. ST could be used to commit torture	☐	☐	☐
3. QUESTIONS OF LEGALITY			
3a. Restrictions on the use/sale/etc. of similar STs	☐	☐	☐
3b. Design elements provoke legal challenges	☐	☐	☐
3c. Citizens cannot determine illegal use of the ST	☐	☐	☐
3d. Illegal use not prevented by design safeguards	☐	☐	☐
3e. Rule of law not respected	☐	☐	☐
3f. Operation of ST can entail discrimination	☐	☐	☐
3g. Design doesn't ensure data protection compliance	☐	☐	☐

4. FINANCIAL COST OF THE ST

- 4a. Costs outweigh the security benefits ☐ ☐ ☐
- 4b. Costs considered too high/excessive ☐ ☐ ☐
- 4c. Published costs not believed ☐ ☐ ☐

5. PUBLIC & END-USER ACCEPTABILITY

- 5a. All design aspects not justified ☐ ☐ ☐
- 5b. Public trust in the ST is lost ☐ ☐ ☐
- 5c. End-user trust in the ST is lost ☐ ☐ ☐
- 5d. ST is not necessary ☐ ☐ ☐
- 5e. Minority groups bear the security burden ☐ ☐ ☐
- 5f. ST represents a disproportionate response ☐ ☐ ☐
- 5g. Public support for ST is conditional ☐ ☐ ☐
- 5h. ST doesn't meet social acceptability standards ☐ ☐ ☐

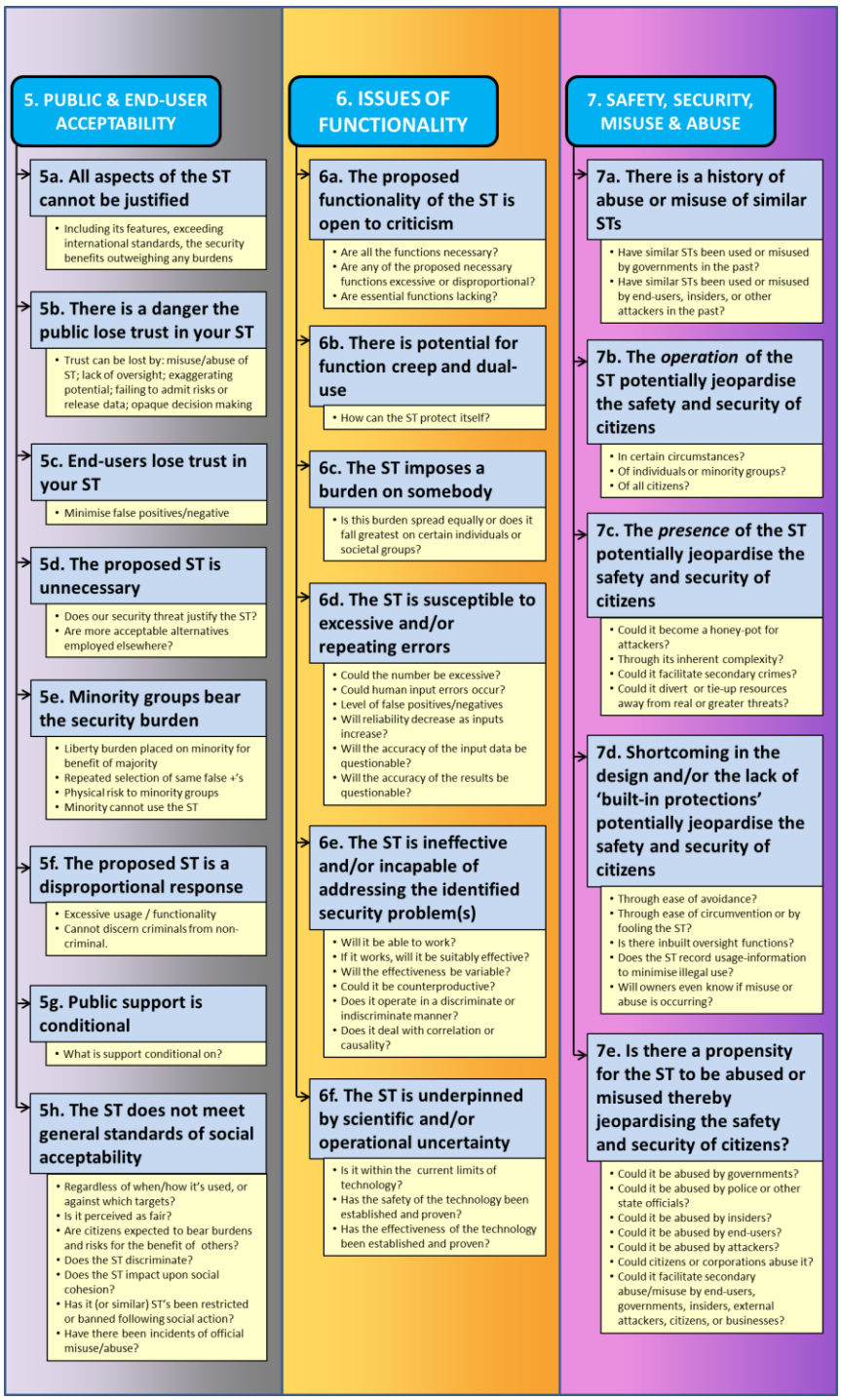
6. ISSUES OF FUNCTIONALITY

- 6a. STs functionality is criticised ☐ ☐ ☐
- 6b. Potential for function creep and dual-use ☐ ☐ ☐
- 6c. The ST imposes a burden on somebody ☐ ☐ ☐
- 6d. ST is susceptible to excessive/repeating errors ☐ ☐ ☐
- 6e. ST ineffectively at address the security problem ☐ ☐ ☐
- 6f. ST underpinned by uncertainty ☐ ☐ ☐

7. SAFETY, SECURITY, MISUSE & ABUSE

- 7a. History of abuse/misuse of similar STs ☐ ☐ ☐
- 7b. STs operation could jeopardise citizens ☐ ☐ ☐
- 7c. STs presence could jeopardise citizens ☐ ☐ ☐
- 7d. Design flaws/omissions jeopardise citizens ☐ ☐ ☐
- 7e. Propensity for the ST to be abused/misued ☐ ☐ ☐


1. PHYSICAL OR MENTAL HARM
1a. Health risks created by the security technology (ST) outweigh the potential security benefits of that technology
• Under the principle of *justification* the benefits should outweigh the harms.
1b. The health risks are unacceptable to the public
• Can the public be convinced that the benefits outweigh the costs?
• Will vulnerable groups that society is protective of be placed at greater risk? (e.g. elderly, pregnant women, mental patients, children, etc.)
1c. The health risk greater for *at-risk* minority groups
• Are there additional health risks for identifiable minority groups within society? (e.g. children, pregnant women, elderly, etc.)
• Are there additional health risks for unidentifiable groups within society? (e.g. asthma, heart conditions, early pregnancy, mental conditions, those on medication, etc.)
1d. The ST causes pain or discomfort
• Including nausea, dizziness, pain, impairment, headaches, etc.
1e. The ST causes physical injuries
• Either permanent or transient injuries
1f. The ST can cause a fatality
• By STs designed not to kill
1g. Designers fail to be open and honest about the health risks of their ST
• Is it too risky to undertake extensive live human testing?
• How would you react if these health risks were identified and published by independent researchers in peer-reviewed journals?
• Any denial of actual health risks is undermined by statements to the contrary from judges and coroners.
2. LIBERTIES & HUMAN RIGHTS
2a. The ST impacts ones right to privacy
• Informational, physical, or image
• Could it have a chilling effect?
2b. The ST fails to respect informational self-determination
• Does it collect / process data without knowledge or consent?
• Can subjects control their data?
2c. The ST impacts ones right to assembly and association
• Could it have a chilling effect?
• Could it prevent enjoyment of this right?
2d. The ST impacts freedom of expression
• Either directly or by a chilling effect?
2e. The ST impacts the ability of a citizen to have a fair trial
• Is their a presumption of innocence?
• Are the ST's decision making processes opaque ?
• Can the subject defend themselves?
• Can 'reasonable suspicion' be determined?
2f. The losses to other human rights outweigh the security benefits
• Is the ST necessary?
• It the ST a proportional response to the problem being addressed?
• Will the negative consequences of the ST be outweighed by the potential benefits?
2g. The ST can be used for torture or inhumane or degrading treatment
• Can/has this or similar security technologies been used as a torture device?
• Could it possibly be used in a manner could be considered inhumane or degrading?
3. QUESTIONS OF LEGALITY
3a. The use, sale, or possession of similar STs restricted or banned?
• Are there specific export bans?
3b. Elements of design or operation are likely to provoke legal challenge
• Previous issues include: images of children; disability, race, sex discrimination; data protection; noise; harassment; etc.
3c. Citizens cannot determine whether the ST is being operated legally
• Opaque decision making process
• Unchallengeable decisions
• Lack if information makes determining legality impossible
3d. illegal use of the ST is not prevented design safeguards
• Are there obligations to prevent unauthorised use?
• Is data protected?
• Is use monitored to prevent unlawful use?
3e. There are not design protections within the ST to ensure compliance with all data protection principles
• Lack of transparency:
➢ no consent to the collection or processing of data
➢ Individuals not informed when personal data is processed
➢ No ability to rectify, delete, or block access to personal data that is incomplete or unlawfully processed
• Lack of legitimate purpose:
➢ Data is processed for non-specific purposes
➢ Data is processed for non-legitimate purposes
• Lack of proportionality:
➢ Excessive collection, processing, or storage of data
➢ Lack of anonymisation
• Failure to protect stored data
• No obligation to protect data or prevent unauthorised use
• Unmonitored use may result in illegal use
3f. The rule of law is not respected
• Is the application of the ST both clear and predictable?
• Will it be applied equally to all?
3g. Operation of the ST can entail discrimination
• Based on : sex, age, religion, religious dress, nationality, ethnicity, disability, etc.
• Does it create arbitrary categories?
4. FINANCIAL COST OF THE ST
4a. The financial costs outweigh the security benefits
• Do the security gains justify the price-tag?
• Can similar gains be achieved more cheaply through other means?
4b. The cost of this technology is considered too high or excessive
• How high will the initial purchasing and set-up costs be?
• How high are the on-going operation & maintenance costs?
4c. The financial figures released into the public domain are not trusted
• Are the benefits and costs believable?
• Have they been independently produced or confirmed?


5. PUBLIC & END-USER ACCEPTABILITY
5a. All aspects of the ST cannot be justified
• Including its features, exceeding international standards, the security benefits outweighing any burdens
5b. There is a danger the public lose trust in your ST
• Trust can be lost by: misuse/abuse of ST; lack of oversight; exaggerating potential; failing to admit risks or release data; opaque decision making
5c. End-users lose trust in your ST
• Minimise false positives/negative
5d. The proposed ST is unnecessary
• Does our security threat justify the ST?
• Are more acceptable alternatives employed elsewhere?
5e. Minority groups bear the security burden
• Liberty burden placed on minority for benefit of majority
• Repeated selection of same false +'s
• Physical risk to minority groups
• Minority cannot use the ST
5f. The proposed ST is a disproportional response
• Excessive usage / functionality
• Cannot discern criminals from non-criminal.
5g. Public support is conditional
• What is support conditional on?
5h. The ST does not meet general standards of social acceptability
• Regardless of when/how it's used, or against which targets?
• Is it perceived as fair?
• Are citizens expected to bear burdens and risks for the benefit of others?
• Does the ST discriminate?
• Does the ST impact upon social cohesion?
• Has it (or similar) ST's been restricted or banned following social action?
• Have there been incidents of official misuse/abuse?
6. ISSUES OF FUNCTIONALITY
6a. The proposed functionality of the ST is open to criticism
• Are all the functions necessary?
• Are any of the proposed necessary functions excessive or disproportional?
• Are essential functions lacking?
6b. There is potential for function creep and dual-use
• How can the ST protect itself?
6c. The ST imposes a burden on somebody
• Is this burden spread equally or does it fall greatest on certain individuals or societal groups?
6d. The ST is susceptible to excessive and/or repeating errors
• Could the number be excessive?
• Could human input errors occur?
• Level of false positives/negatives
• Will reliability decrease as inputs increase?
• Will the accuracy of the input data be questionable?
• Will the accuracy of the results be questionable?
6e. The ST is ineffective and/or incapable of addressing the identified security problem(s)
• Will it be able to work?
• If it works, will it be suitably effective?
• Will the effectiveness be variable?
• Could it be counterproductive?
• Does it operate in a discriminate or indiscriminate manner?
• Does it deal with correlation or causality?
6f. The ST is underpinned by scientific and/or operational uncertainty
• Is it within the current limits of technology?
• Has the safety of the technology been established and proven?
• Has the effectiveness of the technology been established and proven?
7. SAFETY, SECURITY, MISUSE & ABUSE
7a. There is a history of abuse or misuse of similar STs
• Have similar STs been used or misused by governments in the past?
• Have similar STs been used or misused by end-users, insiders, or other attackers in the past?
7b. The *operation* of the ST potentially jeopardise the safety and security of citizens
• In certain circumstances?
• Of individuals or minority groups?
• Of all citizens?
7c. The *presence* of the ST potentially jeopardise the safety and security of citizens
• Could it become a honey-pot for attackers?
• Through its inherent complexity?
• Could it facilitate secondary crimes?
• Could it divert or tie-up resources away from real or greater threats?
7d. Shortcoming in the design and/or the lack of 'built-in protections' potentially jeopardise the safety and security of citizens
• Through ease of avoidance?
• Through ease of circumvention or by fooling the ST?
• Is there inbuilt oversight functions?
• Does the ST record usage-information to minimise illegal use?
• Will owners even know if misuse or abuse is occurring?
7e. Is there a propensity for the ST to be abused or misused thereby jeopardising the safety and security of citizens?
• Could it be abused by governments?
• Could it be abused by police or other state officials?
• Could it be abused by insiders?
• Could it be abused by end-users?
• Could it be abused by attackers?
• Could citizens or corporations abuse it?
• Could it facilitate secondary abuse/misuse by end-users, governments, insiders, external attackers, citizens, or businesses?